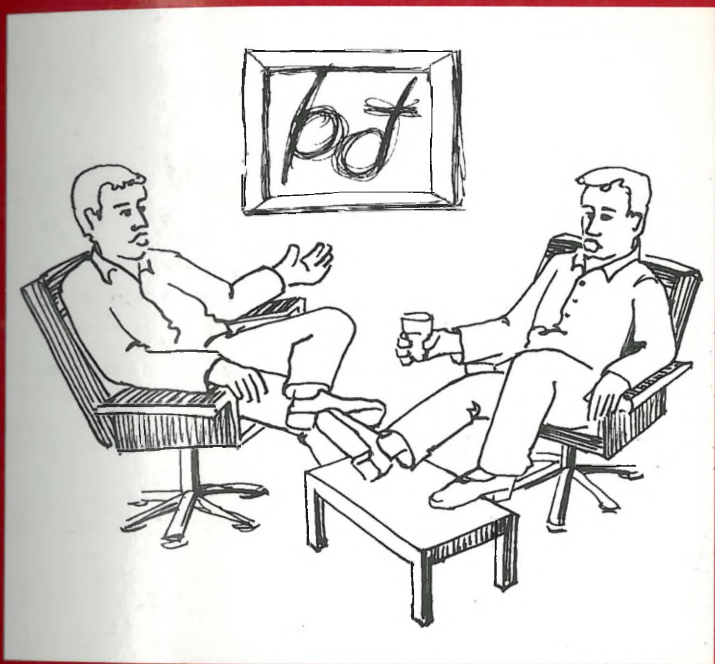
Mirjam Ernestus

Voice Assimilation and Segment Reduction in Casual Dutch

A Corpus-Based Study of the Phonology-Phonetics Interface



HIL

HOLLAND INSTITUTE OF GENERATIVE LINGUISTICS

LOT

Netherlands Graduate School of Linguistics

Landelijke Onderzoekschool Taalwetenschap a

Voice assimilation and segment reduction in casual Dutch

a corpus-based study of the phonology-phonetics interface

Published by

LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6006
fax: +31 30 253 6000
e-mail: lot@let.uu.nl
http://www.let.uu.nl/LOT/

ISBN 90-76864-02-0
NUGI 941

Cover illustration by Van Herwijnen Design.

Printed in the Netherlands

Acknowledgements

I am glad the moment has come at which I can finally do justice to all those who helped me in the past five and half years to do my research and write this dissertation.

First of all I would like to thank the Netherlands Organization for Scientific Research (NWO) and the Holland Institute of Generative Linguistics (HIL) for providing me with money to attend three linguistic summer schools and conferences in Berlin and York.

Then, I would like to thank Geert Booij, who has been my primary supervisor and promotor. Over the years I have come to esteem him higher and higher, in particular for his professionalism. He proved to be a quick and careful reader, did his very best to create the optimal circumstances for me to write my dissertation, and gave me complete freedom in choosing my topic.

When I had decided to create a corpus myself and to transcribe it with the help of a speech analysis software package, I received a lot of help from the members of the Institute of Phonetics of the University of Amsterdam. I am grateful that they allowed me to use their sound proof room and their computer equipment. In particular, I would like to thank Ton Wempe for his never ending assistance.

Although working at the phonetic lab of the University of Amsterdam was convenient, I preferred to work "at home" at my office at the Vrije Universiteit Amsterdam. I am very much indebted to Paul Boersma, who was willing to spend many days installing his speech analysis software package *Praat* on the more or less incomplete UNIX computer of the Faculty of Lettres of the VU. Above all, however, I would like to thank him for the discussions we had during the last four years, which strongly influenced my linguistic ideas.

Creating a speech corpus is possible only if there are subjects willing to fill tapes with their speech. I would like to cordially thank all subjects who participated in one of the experiments. In addition, I would like to thank the six speakers who judged the speech taped for the actual corpus.

The major part of the corpus was orthographically transcribed by Rosalie van de

Sandt. I am very grateful that she offered to do this job for me. Above all, however, I would like to thank her for her friendship, and for agreeing to be my "paranimf".

Thousands of stops and hundreds of stretches of speech from the corpus were transcribed by Boudewijn van den Berg, Marian Klamer, and Piet van Reenen, for which I am very grateful. Special thanks are also due to Boudewijn because of his enthusiasm and confidence in me. I am indebted to Piet also for reading all chapters of this book when they were still in an embryonic stage. His comments proved to be very useful.

When the stops had been transcribed and the measurements taken, I had to process the data. Evert Wattel helped me enormously with this task. He drew hundreds of logarithmic curves (which I finally decided not to incorporate) on the basis of my data, developed the method to separate clusters in two-dimensional plots which is described in Appendix F, and produced nearly all the pictures in this dissertation. I am very grateful for the time he spent on behalf of me, and admire his cheerfulness and the ease with which he adapted the pictures and method whenever I wanted.

During the first two years of my project I had the luxury of having an office nearly all to myself. Then I had to move in with someone, and chose Onno Huber. Onno proved to be the ideal office mate. He was with me in every stage of the project, and read my dissertation before I offered it for final judgement to my supervisors. Thank you Onno.

After four years of working on the project, Geert Booij and I realized that a copromotor was needed, and that Toni Rietveld was the obvious person. I am glad that Toni accepted the job. His comments, in particular, on the statistical part of this dissertation proved to be very useful. I am grateful for his enthusiasm and confidence that this dissertation would be interesting at those moments when I was not so sure of it myself at all.

In addition to Piet van Reenen and Onno Huber, there were others who in the course of time have read (parts of) my dissertation and provided me with valuable comments. I would like to thank Carlos Gussenhoven, Vincent van Heuven, Patricia Keating, Evert Wattel, Leo Wetzels, Sieb Nooteboom, Iman Slis, and Ruben van de Vijver.

When the job was done, and the book could be sent to the committee, Bettelou Los was so kind to fish the English mistakes from the text. Within a week she read the whole book and gave many valuable suggestions on how to make the text more fluent, and more intelligible for linguists other than phonologists and phoneticians. She deserves many thanks for that.

Just before the book was published, Hans van Herwijnen drew the picture on the front of this book. I am happy he was willing to draw the picture as I wanted it to be. Thank you.

During these five and a half years at the Vrije Universiteit, I never had a chance to feel alone, as there was a whole bunch of Ph.D. students and post-docs who made my lunches agreeable, and helped me with all kinds of little things. I would like to thank Hilde van Aken, Maaike Beliën, Frank Beths, Ronny Boogaart (who looked for English mistakes in my LIN-article), Hanny Demeersseman, Wilma Elsing, Nelleke Goudswaard (who has been my, very appreciated, second office mate since last December), Rick de Graaff, Marjan Huisman, Thera de Jong, Karolien Kamma, Marian Klamer, Gerson Kuiper, Nienke Lammersen-van Deursen, Claartje Levelt, Machiel Limburg (who spent some hours working on my bibliography), Marga Petter, Cristina Pumplun, Margit Rem, Ruben van de Vijver, and Astrid Wijnands.

In addition, I would like to thank Audra Dainora with whom I spent 14 weeks of linguistic summer schools. Thanks to her these schools were more like holidays. I would like to thank her also for coming to the Netherlands a few days earlier than originally planned in order to be my "paranimf".

Tenslotte wil ik iedereen bedanken die er rekening mee heeft gehouden dat ik door "dat proefschrift" niet altijd zoveel tijd aan hen (of aan onze zwemvereniging) kon besteden als ik wel wilde. Vooral mijn ouders wil ik hartelijk bedanken. Ik waardeer zeer de manier waarop zij mij tijdens mijn studie en AIO-periode hebben bijgestaan.

Contents

Part IV: A rough survey of phoneme realizations

6 Absent segments and reduced vowels

10 Geminate alveolar stops

Part VI: Final Remarks

11 Conclusions

Appendices

C Instructions for the role-play

D Fragments of the corpus

E Judgements of the recorded speech

F Separating two clusters in two-dimensional plots

Abbreviations and symbols

dim.	diminutive
fem.	feminine
inf.	infinitive
pres.	present
pron.	pronoun
ps.	person
rel.	relative
sg.	singular
F	foot
PrWd	prosodic word
V	vowel
(	left boundary of prosodic domain
)	right boundary of prosodic domain
σ	syllable

Part I

Preliminaries

1 Introduction

1.1 Casual speech

Whereas the pronunciation of careful Modern Standard Dutch has been investigated extensively, witness the publications of Zwaardemaker & Eijkman (1928), Blancquaert (1934), Jongenburger & van Heuven (1993), Booij (1995), and many others, little is known of the pronunciation of casual Dutch. The present study will try to fill this gap: it aims to present new data, and to provide a better understanding of the mechanisms behind casual Dutch.

Casual speech, which will be defined here as the every-day speech used in informal situations in which no special attention is paid to expression (cf. Labov 1972: 86), is in general quite underinvestigated. A survey of works on casual English and German, for instance, reveals only a small number of studies (e.g. Zwicky 1972; Brown 1977; Dalby 1984; Kaisse 1985; Kohler 1990; Keating 1998), and the casual variants of many other languages are still completely unexplored. There are several explanations for this.

One explanation is that obtaining data on casual speech is difficult. Speakers do not have valid intuitions on all aspects of the phonetic realizations of casual speech, which is why studies of these realizations cannot be exclusively based on speakers' intuitions, but have to be based at least partially on actual speech. Observing realizations in casual speech is difficult since this style of speech is generally fast.

Another explanation is the fact that casual speech has long been believed not to deserve to be a research field in its own right. Both de Saussure (1916) and Chomsky & Halle (1968) have claimed that the study of language should be concentrated on the mental system which is responsible for all human behaviour with respect to language, including linguistic intuitions, speech errors, and language change. Actual speech, and especially casual speech, would be governed not only by this mental system, but also by the articulatory and perceptual restrictions of human speakers. As a consequence, any regularities found in casual speech would not necessarily be the product of the mental system, and the mental system should not be investigated on the basis of data from casual speech. It is only recently that the mental language system is assumed to govern all regularities in speech (§2.4.1), and that consequently casual speech provides relevant information on this system. This view entails that casual speech requires an investigation, since it may display regularities unlike any found in other types of language behaviour.

The final reason why so few studies have been devoted to casual speech may be that such studies were regarded as having no social relevance. Orthographies and textbooks for foreign language learners, for instance, were designed on the basis of studies of careful speech. Here, too, things have changed: studying casual speech has now become vital to the development of automatic speech recognition, to name just one application.

The few studies on casual speech to date show that many types of assimilation and elision which are not permitted in the careful variant of a language apply frequently in the casual variant. Keating (1998), for instance, found that words which have only one form in careful English may have as many as fourteen forms, i.e. a full form and thirteen reduced forms, in casual English. Function words in particular often surface as highly reduced forms. An example is the German function word *einen* 'inflected definite article', which is generally realized as [ʔaɪnən] in careful speech, but as [n] in casual speech (Kohler 1990, 1998).

Most analyses which try to account for the observations of casual speech claim that the differences between careful and casual speech are mainly due to the difference in importance that speakers attach to ease of articulation and ease of perception in the two speech styles. When people speak carefully, they try to be as intelligible as possible. This implies that they attach great importance to ease of perception, and articulate as clearly as possible, sacrificing ease of articulation. When people speak casually, on the other hand, they attach less importance to ease of perception. They still want to be intelligible, but ease of perception is sacrificed to ease of articulation to some extent. This idea is incorporated in Lindblom's (1990) H&H theory, which states that the compromise between ease of perception and ease of articulation varies along a continuum of hyperspeech and hypospeech. The compromise is speaker specific, as there are speakers who speak carefully under all circumstances (cf. Zwicky 1972: 607).

Increase in ease of articulation, that is reduction of articulation effort, implies reduction of the sizes of the articulatory gestures. This is why it can result, for instance, in the realization of a stop as a fricative. The gesture that has to form the constriction in the vocal tract for the realization of the stop is then not finished, and the resulting constriction is not complete, but has a degree characteristic for fricatives.

In addition, increase in ease of articulation probably implies the less exact timing of the articulatory gestures. Instead of waiting for the exactly defined moment to start a certain articulatory gesture and then instantly realizing this gesture, a speaker may choose to start the gesture somewhat earlier, and realize it at a lower speed. This generally results in the coarticulation of segments.

When a segment is co-articulated completely with other segments, it can be perceptually absent. The [t] of the phrase *perfect memory*, for instance, can be perceptually absent because its realization coincides with the constrictions of the preceding [k] and the following [m] (Browman & Goldstein 1990a: 364). When segments are realized simultaneously only partially, the result may appear to be feature assimilation. Vowel assimilation across Prosodic Word-boundaries in Igbo is an instance of regressive assimilation which results from the retiming of gestures (Zsiga 1997).

The existing analyses of casual speech call for an in-depth investigation into the mechanisms responsible for realizations in casual speech, as does the recent view that casual speech reveals the properties of the mental system, and that casual speech is socially relevant. In this book, we will provide new data, and discuss whether these data suggest an important influence of the speaker's natural tendency to reduce articulatory effort, or the influences of other factors. By "reduction of articulatory effort" we mean the reduction of the sizes of articulatory gestures, and the increase in their temporal overlap. We will concentrate on those varieties of Dutch which are accepted as standard by the great majority of speakers of Dutch, that is, Standard Dutch.

1.2 Topic of the study

The exact questions which will be addressed in this study are:

1. In which contexts
    - are full vowels realized as schwas;
    - do segments tend to be absent;
    - are obstruents realized as voiced or voiceless;

   in casual Modern Standard Dutch?

2. What is the role of the speaker's natural tendency to reduce articulatory effort in the casual realizations?

3. Which other factors may be responsible for the casual realizations?

This study consists of a rough survey and an in-depth investigation. The former will deal with the realization of full vowels as schwas, and the perceptual absence of segments. We will argue that they are the possible result, in the majority of cases, of the speaker's natural tendency to reduce articulatory effort, and are partly a matter of the lexicon.

The in-depth study, on the other hand, concentrates on the realization of obstruents as voiced or voiceless. This subject has been chosen for several reasons. First, we will argue that coda obstruents and word-final obstruents are realized as voiced or voiceless depending on which realization requires the least articulatory effort, and the realization of these segments is therefore very relevant to the theme of this study. Second, obstruents have been a favourite subject of many phonological studies of Dutch, which indicates that the realization of obstruents in this language is generally assumed to be interesting and relevant to phonological theory. Finally, there is the benefit of being able to build on the results of previous studies, which allows us to discuss detailed analyses. Since some data are still lacking, and previous analyses are unable to explain all available data, a new study is necessary.

We will propose an analysis of the realizations of obstruents as voiced or voiceless in casual Dutch which assumes that word- and syllable-final obstruents are not specified for [voice] in the output of phonology, and that it is the phonetic component that determines whether they are realized as voiced or voiceless. They are realized as voiced when a voiced realization requires less articulatory effort, and as voiceless when it is the voiceless realization that takes the least articulatory effort. Because the coda obstruents are assumed to be unspecified for [voice], the voiceless realization of some types of underlyingly voiced segments which follow these obstruents cannot be considered to be instances of phonological [voice]-assimilation, as in previous analyses. We will claim that underlyingly voiced fricatives and the initial /d/s of some function words are realized as voiceless after coda obstruents because obstruents in clusters are preferably not voiced, and these obstruents are allowed to surface with [voice]-specifications which differ from their underlying ones in order to be voiceless in clusters. The initial stop of the regular past-tense morpheme *-te/de*, on the other hand, is realized as voiceless after underlyingly voiceless coda obstruents because it takes over the underlying [voice]-

specifications of preceding obstruents that must be unspecified for [voice] in the output of phonology. Overall, the analysis incorporates several hypotheses with respect to the realization of intervocalic obstruents, and word-initial /d/s. Some of them, including one on the influence of the lexicon, will be tested on the basis of casual speech in this study, and found to be correct.

The rough survey and the in-depth investigation constitute the core of this study. In addition, much attention will be paid to methodology, since the methods by which the relevant data are obtained and processed may not be obvious. We will argue that realizations in casual speech have to be studied on the basis of recorded speech from the spontaneous conversations of people who are acquaintances of each other. A corpus containing such speech was built specifically for this study. In addition, we will extensively motivate the types of words and word-combinations which were chosen as the basis for the investigations, and deal with the classification of obstruents as either voiced or voiceless.

1.3 Outline of the book

The remainder of this book has been divided into five parts. Part II consists of Chapters 2 and 3, which will describe the theoretical background of the study. Chapter 2 will focus on relevant properties of the lexicon, and the phonological and phonetic components. Chapter 3 will present relevant parts of Dutch phonology, including previous analyses of the realizations of obstruents as voiced or voiceless.

Part III consists of Chapters 4 and 5, and will discuss the type of data forming the basis of the investigations. In Chapter 4, we will argue that the investigations must be neither based on linguistic intuitions nor on recordings of read aloud speech, but on corpora of spontaneous conversations. The corpus that has been built for the purpose of the present study will be described in detail in Chapter 5.

Part IV consists of Chapter 6, which will present the rough survey of the perceptual absence of segments and the realization of vowels as schwas in casual Standard Dutch. This chapter will reveal new data, and provide clues to the relevance of the speaker's natural tendency to reduce articulatory effort and the role of the lexicon with respect to realizations in casual speech.

Part V will present the in-depth study into the realization of obstruents as voiced or voiceless in Dutch. We will argue that previous analyses of the realizations of obstruents are not satisfactory, and propose an alternative in Chapter 7. Four hypotheses incorporated in this alternative analysis will be tested in the remainder of Part V. Chapter 8 will extensively describe the methodology of the testing procedure, while Chapters 9 and 10 will test the hypotheses on the basis of the realizations of intervocalic stops and geminate coronal stops respectively.

Finally, Part VI consists of Chapter 11. It will relate the results obtained in the different parts of this book, and recapitulate the most important conclusions.

Part II

Theoretical background

2 The lexicon, phonology, and phonetics

2.1 Introduction

This book adopts an organization of the relevant part of the grammar as sketched in Figure 2.1.

Underlying form stored in the lexicon

⇓

$\boxed{\text{Lexical level of Phonology}}$

⇓

$\boxed{\text{Post-lexical level of Phonology}}$

⇓

Phonological form

⇓

$\boxed{\text{Phonetics}}$

⇓

Phonetic form (Acoustic form)

Figure 2.1 The relevant parts of the grammar (represented by boxes), and their inputs and outputs.

It assumes that generally all surface forms of a morpheme can be computed on the basis of one string of phonemes, which is the underlying form. This form is independent of factors such as the rate of speech, speech style, gender of the speaker, etc. It is stored in the lexicon, and forms the input to the lexical level of phonology.

The output of Lexical Phonology forms the input to Post-lexical Phonology, whose outputs are strings which will be referred to as "phonological forms". These forms contain phonemes, like the underlying forms, and are similarly independent of factors such as the characteristics of the speaker, and the speech rate. They are generally more narrowly related to the actual realizations of the strings than the underlying representations, and form the input to the phonetic component.

The phonetic component has as its output the actual realizations, which will sometimes be called the "phonetic forms". They can be characterized, for instance, in terms of the articulatory gestures involved, and the acoustic consequences of these gestures. The properties of phonetic forms are influenced by characteristics of the speaker, speech style, speech rate, etc. The term "acoustic form" refers to the audible characteristics of a phonetic form.

This chapter identifies and discusses in detail all relevant assumptions with respect to the lexicon (§2.2), phonology (§2.3) and phonetics (§2.4).

2.2 The lexicon

2.2.1 Types of stored units

Following Chomsky & Halle (1968), the present study does not assume that all morphologically complex words or string of words which may be uttered are stored in the mental lexicon. A subset of them is always computed, or accessed, from the underlying forms of their parts. This assumption explains the ability of speakers to produce and understand words and sentences they have never heard before.

The lexicon contains units whose characteristics are unpredictable: it contains the representations of all morphologically simplex words and morphologically complex strings with unpredictable semantic, morphological, phonological, or syntactic characteristics. The lexicon of a speaker of English, for instance, contains representations of *book* (a simplex word), *kept* (a morphologically complex word with an unpredictable appearance), *cupbearer* (a morphologically complex word with an unpredictable meaning), and *need badly* (a string of words with an unpredictable meaning). The stored representations include information on the semantic, syntactic, morphological, and phonological characteristics of the units (see e.g. Levelt 1989: 182).

The capacity of human memory is very large, and storage may be less costly than parsing. The lexicon therefore possibly contains not only representations of simplex and irregular complex units, but also of several regular complex units (contra Chomsky & Halle 1968). The results of experiments suggest that this is indeed the case: the lexicon appears to contain regular units with high frequencies of occurrence. For instance, Stemberger & MacWhinney (1988) found that the number of errors made in the production of regularly inflected forms is significantly higher for low frequency forms than for high frequency forms. If smaller numbers of errors indicate less computation or parsing, which seems a reasonable assumption, these findings suggest that highly frequent forms tend to be retrieved as units from the lexicon, whereas the low frequency ones are generally computed from their parts. The influence of frequency of occurrence on the storage of regular complex units is not surprising, since storing complex units saves a lot of processing power especially if the units are extremely frequent. Apparently, the savings in processing costs counterbalance the storage costs only if the units are highly frequent, and consequently it is only the storage of highly frequent strings that is economical.

It is possible that units are only stored in the lexicon for a limited period of time. This may be the case when strings that generally have a low frequency keep cropping up in a particular conversation. These strings are probably stored for the duration of that conversation, since their storage introduces savings in the processing costs which counterbalance the increase in the memory load at that moment.

In conclusion, the lexicon minimally contains the representations of units which are morphologically simplex, or have unpredictable characteristics, or are highly frequent. Section 2.2.2 will deal with the organization of the lexicon, while section 2.2.3 will discuss the influence of the lexicon on the realization of words and phonemes.

2.2.2 Organization of the lexicon

Several models have been proposed for the organization of the lexicon, the most recent of which generally assume that the way in which an item is stored in the lexicon depends on its frequency of occurrence (see e.g. Garnham 1985: 46-53). This assumption is based on two observations. The first is that when people are asked to read aloud words, they start reading earlier if the words are of high frequency of occurrence (Solomon & Howes 1951ab; Forster & Chambers 1973). The second observation is that when people have to decide whether or not a given string of letters exists as a word in the language, they react the fastest when they are presented with an existing word of high frequency (Rubenstein et al. 1970). The speed with which a unit is accessed apparently depends on its frequency of occurrence, which implies that the way it is stored depends on frequency as well.

Probably, the segments of a unit are not all stored in the same way, and there may be a difference in representation between its initial, final, and stressed segments and its other segments. When people have words on the tips of their tongues, they often remember the first and last segments of these words, and when they mix up words, these words often share their first and final segments and stressed syllables. These segments are apparently prominent in storage (see e.g. Aitchison 1987: 121), and highly relevant for the recognition of units.

There is no consensus on the lexical forms of units containing several words. Some linguists, including Booij (1985), assume that the lexical form of such a unit contains exactly the same segments as the sum of the lexical forms of its parts. This view implies, for instance, that if the word combinations *heb ik*, consisting of *heb* /hɛb/ 'have' and *ik* /ɪk/ 'I', and *weet ik* consisting of *weet* /ʋet/ 'know'and *ik* /ɪk/ 'I' are stored in the lexicon, they have the lexical forms /hɛbɪk/ and /ʋetɪk/ respectively. Other linguists, e.g. Bybee and Scheibman (Bybee 1995, 1996; Bybee & Scheibman 1999), claim that the lexical form of a string of words represents its usual realization. Thus, if the strings *heb ik* and *weet ik* are generally realized as [hɛpɪk] and [ʋedɪk], and are stored as units in the lexicon, their lexical forms are /hɛpɪk/ and /ʋedɪk/.

2.2.3 Lexical effects on realization

Units which are stored in the lexicon may differ in appearance from those which are not. The acoustical length of a segment, for instance, can be partly determined by its presence or absence in the lexicon. This appears e.g. from Losiewicz (1992), who shows that the coronal stop in English is shorter if it belongs to a monomorphemic word or is the past-tense morpheme of a verb form of high frequency than if it is the past-tense morpheme of a verb form of low frequency. Since monomorphemic words and highly frequent complex words are probably stored in the lexicon, these data strongly suggest that segments which are present in the lexicon are acoustically shorter than segments which are absent.

Stored and non-stored units may also differ in their prosodic structure. This is evident, for instance, from the prosodic structures of morphological compounds in Dutch, since some of the stored compounds form single prosodic words whereas the non-stored compounds all form several ones (§3.6 for relevant assumptions on prosodic structure). The difference between the words *aardappel* /ardɑpəl/ 'potato' and *marsepeinappel* /mɑrsəpɛin-ɑpəl/ 'an apple made of marzipan' is a case in point. The word *aardappel* is formally a compound consisting of *aard* 'earth' and *appel* 'apple'. It is highly frequent, since potatoes are a traditional staple of Dutch dinners. Being highly frequent, *aardappel* is stored in the lexicon, and people can access it as a unit, as is evident from the fact that they no longer regard it as a kind of apple, but as a completely separate type of vegetable. The compound *marsepeinappel* consists of the parts *marsepein* 'marzipan' and *appel* 'apple'. It is not stored in the lexicon, as it is of extremely low frequency of occurrence, and has no unpredictable characteristics. Native speakers claim that the second syllable of *aardappel* starts with [dɑ]. This implies that the two parts of the word form one domain of syllabification, i.e. one prosodic word (§2.3.4). The word *marsepeinappel*, in contrast, forms several prosodic words. Consulted native speakers claim that its fourth syllable starts with [ɑ], i.e. that it is an onsetless syllable preceded by a consonant. Since Dutch generally respects the Maximal Onset Principle (§3.6), the onsetless syllable must be the result of the presence of a prosodic word boundary before *appel*, which forms a boundary to the syllabification domain. The stored compound *aardappel* and the non-stored compound *marsepeinappel* differ in their prosodic structure.

Finally, Booij (1985) claims that when a string is accessed as a unit in the lexicon, its phonological form can differ in its segments from the phonological form that is obtained when the string is computed from its parts. These differences are not assumed to be present yet in the lexical forms, since the lexical forms of strings are assumed to contain the same segments as the sum of the lexical representations of their parts (§2.2.2). Booij assumes that the differences are introduced by lexical phonological rules (see §2.3.3 for the difference between lexical and post-lexical phonology). When a string is accessed as a unit, it is subject to all lexical rules as a unit. When it is computed from its parts, the parts undergo these rules separately, and therefore can be affected by rules which do not affect the units, or, conversely, can be insensitive to rules which do affect the unit. The string *heb ik*, consisting of the parts /hɛb/ 'have' and /ɪk/ 'I', is a case in point. Booij claims that when this string is computed from its parts, its phonological form is [hɛpɪk]. The word *heb* is dealt with in isolation at the lexical level, and the /b/ is in coda position when Final Devoicing applies (§3.4.2). Final Devoicing consequently changes the /b/ into [p]. When *heb ik* is accessed as a unit, its phonological form is [hɛbɪk]. It forms a unit at the lexical level, and the /b/ is in onset position, and is not affected by Final Devoicing.

In summary, there is a difference between units which are and which are not stored in the lexicon. This implies that the lexicon influences the realization of the units that it contains. It is as yet unknown in exactly which respects the appearance of stored and non-stored units may differ, and it is tempting to ascribe all unexpected differences between groups of high and low frequency regular complex units to the fact that the former group is represented in the lexicon, whereas the latter group probably is not. This is naturally only acceptable if the relevant differences also exist between the low frequency regular complex units and the units which are certainly stored in the lexicon, or if they can be explained by the influence of the lexicon.

The lexicon influences the realizations of words in yet another way. As mentioned in section 2.2.2, not all words and segments of words are stored in the same way. The way a word is stored depends on its frequency of occurrence, and there is a difference in prominence of storage between the initial, final, and stressed segments of a unit and its other segments. This influences pronunciation.

The influence of frequency of occurrence on pronunciation is evident from the observation that words tend to surface in more reduced forms if they are of higher frequency (see e.g. van Bergem 1995; Booij 1995: 130). Since listeners recognize units of high frequency more easily than units of lower frequency, a speaker can hypo-articulate the highly frequent units to a larger extent without running the risk of being misunderstood. The observation therefore strongly suggests that speakers satisfy their tendency to reduce articulatory effort only if the reduction of effort does not seriously increase the risk of being misunderstood by the listener, which is intuitively correct.

The influence of the prominence of storage is suggested, for instance, by the flapping of alveolar stops in English, since flapping is restricted to word-medial stops in unstressed syllables, i.e. to stops which are not prominent in storage. Moreover, the influence of prominence of storage can be gauged from the size of the glottal spreading gesture for aspiration in English, which is influenced by stress and the position of the segment in the word (see Browman & Goldstein 1992a: 168, 170, and the references cited there). In general, the segments which are prominent in storage appear to be realized with greater articulatory effort. This is not surprising, since these segments are more important for recognition: if they are hypo-articulated, the chances are that the relevant unit will not be recognized by the listener.

2.2.4 Summary

The lexicon is assumed to contain at least two types of units: units with unpredictable characteristics and units of high frequency. The stored units differ in accessibility, with those that are highly frequent being assessed more easily than those that are of low frequency. The units that are not stored cannot be accessed in the lexicon, and are computed from their parts.

These assumptions are summarized in Table 2.1. It should be noted that, contrary to what is suggested in this table, frequency of occurrence is a continuum and has a gradual influence on storage, and the speed of access.

Table 2.1 The presence/absence and accessibility of units in the lexicon, as a function of their type, and frequency of occurrence.

Type of unit	Frequency	Stored in the lexicon	Access in the lexicon
Regular complex	low	No	Impossible
	high	Yes	Easy
Simplex and irregular complex	low	Yes	Not easy
	high	Yes	Easy

Realization is influenced by the lexicon. There may be differences between units that are generally accessed as a whole, and units which are generally composed from their parts. Moreover, words which are more easily accessible in the lexicon tend to be hypo-articulated to a greater extent. Finally, segments which are prominent in storage, i.e. the initial and final segments of units as well as the segments of stressed syllables, are realized with greater articulatory effort than other segments.

2.3 Phonology

2.3.1 Introduction

When units are realized, they first enter the phonological component (§2.1), which is assumed in the present study to consist of a lexical and a post-lexical level. The outputs of both levels are determined by the interaction of phonological constraints. Section 2.3.2 describes the theory of constraint interaction adopted here, i.e. Optimality Theory, while section 2.3.3 motivates the assumption of a lexical and a post-lexical level.

The inputs and outputs of phonology are assumed to consist of features ordered on auto-segmental tiers (Goldsmith 1990), with the features forming one segment being linked to the same time slot (cf. McCarthy 1981; Clements & Keyser 1983;

Hayes 1986, among others). Moreover, they are assumed to have prosodic structures, of which the characteristics are described in section 2.3.4.

2.3.2 Optimality Theory

Optimality Theory (OT) has been developed in Prince & Smolensky (1993), McCarthy & Prince (1993ab, 1995), and related work. It assumes that there is a function *Gen* (Generator), which creates an infinite number of candidates for the output of (the different levels of) phonology. In addition, it assumes that there are universal constraints on outputs. There is, for instance, a constraint stating that segments in the output are identical to the corresponding segments in the input, and a constraint stating that obstruents are voiceless. Constraints can be conflicting, like the two constraints just mentioned. This means that outputs cannot obey all constraints, and it is assumed that the constraints are ranked according to their significance. This ranking is language specific. The function *Eval* (Evaluator) determines which of the output candidates best satisfies a constraint ranking given the input at issue. The evaluation is performed recursively, that is, the output candidates are evaluated per constraint, starting with the constraint that is most highly ranked. If a candidate violates a constraint more often than another candidate, it is removed from the set of possible output candidates. The evaluation stops as soon as only one candidate is left. This optimal candidate is the actual output.

The evaluation of the output candidates is visualized with tableaux, an example of which is Tableau 2.1. For reasons of clarity, the constraints in this tableau do not have their generally accepted names, but names which more clearly indicate their natures.

Each OT tableau shows the relevant input in the left-most cell of the upper row (/bɑkt/ in Tableau 2.1). The constraints which are relevant to the part of the evaluation at issue are listed in the remaining cells of this upper row. They are listed in order of decreasing relevance to the language under consideration. Relevant output candidates are listed in the cells of the first column. Stars in the other cells indicate which candidates violate which constraints, with every star indicating one violation. Tableau 2.1 shows, for instance, that the output candidate $(\text{bɑkt})_\sigma$ violates the constraint NO CODA OBSTRUENTS twice, and that candidate $(\text{bɑk})_\sigma(\text{tɑ})_\sigma$ violates both NO CODA OBSTRUENTS and NO INSERTION once. The first violation which makes a candidate non-optimal is marked with an exclamation mark. Hence, the exclamation marks in Tableau 2.1 indicate that candidates $(\text{bɑkt})_\sigma$ and $(\text{bɑk})_\sigma(\text{tɑ})_\sigma$ are non-optimal because they violate NO CODA OBSTRUENTS at least once, that $(\text{bɑ})_\sigma(\text{kɑ})_\sigma(\text{tɑ})_\sigma$ is non-optimal because it violates NO INSERTION, and finally that

$(\text{ba})_\sigma$ is out because it violates NO DELETION twice. The optimal output candidate is marked with a right-pointing finger. Output candidate $(\text{baŋ})_\sigma$ is optimal in Tableau 2.1, because there is no other candidate which completely satisfies NO CODA OBSTRUENTS and NO INSERTION while violating NO DELETION maximally once. Constraints which are not ranked with respect to each other, or of which the ranking is unknown, are separated in a tableau by a dotted, instead of a solid, line (see e.g. Tableau 3.1 in §3.4.2), or are listed in the same cell (see e.g. Tableau 7.2 in §7.5.2.1).

Tableau 2.1 An example of candidate selection.

/bakt/	NO CODA OBSTRUENTS	NO INSERTION	NO DELETION
$(\text{bakt})_\sigma$	*!*		
$(\text{bak})_\sigma(\text{ta})_\sigma$	*!	*	
$(\text{ba})_\sigma(\text{ka})_\sigma(\text{ta})_\sigma$		*!*	
$(\text{ba})_\sigma$			**!
☞ $(\text{baŋ})_\sigma$			*

There are roughly two types of constraints. There are faithfulness constraints, which state that segments (and features) of the output are identical to those of the input (or another form). In addition, there are wellformedness constraints, which state that the output has unmarked properties, or does not have marked properties. Since realizations, and therefore probably also phonological representations, deviate from the corresponding underlying representations more in casual speech than in careful speech, faithfulness constraints are assumed to be ranked more highly in more formal styles of speech (van Oostendorp 1997).

For convenience, we will sometimes refer to the fact that a phonological form differs in a certain respect from the corresponding underlying form as the result of a "process".

2.3.3 The lexical / post-lexical distinction

The present study assumes that phonology consists of a lexical and a post-lexical level. At the lexical level, words are dealt with in isolation, whereas at the post-lexical level context may play a role. All analyses which account for assimilation across word-boundaries are of course forced to take context into account, and therefore to assume a post-lexical level. We also assume a lexical level, because several linguistic facts can be well accounted for under the assumption that certain generalizations hold only for this level.

For instance, the quality with which underlying /e/ is realized in French can be well explained in analyses adopting a lexical level (Booij 1984: 199). This vowel generally surfaces as [e] in open syllables, and as [ɛ] in closed syllables. This is illustrated in (1ab). The correspondent of /e/ is in an open syllable in the phonetic form of *premier* /prəmjer/ (see 1a), and in a closed syllable in *première* /prəmjer+ə/ (see 1b), since the last segments of these underlying forms do not surface in the actual output. The vowel is realized as [e] in *premier*, and as [ɛ] in *première*.

(1) The realization of /e/ in open and closed syllables in French. The dot in the phonological/phonetic form indicates a syllable boundary.

a.	*premier*	/prəmjer/	'first.masc.'	[prəmje]
b.	*première*	/prəmjer+ə/	'first-fem.'	[prəmjɛr]
c.	*première amie*	/prəmjer+ə ami/	'first-fem. friend'	[prəmjɛ·rami]

Contrary to the generalization that /e/ surfaces as [e] in open syllables, it surfaces as [ɛ] in the phrase /prəmjer+ə ami/ [prəmjɛ·rami] (1c). The vowel is part of an open syllable in the surface form of this phrase, but is realized with the quality of /e/ in a closed syllable. If it is assumed that the final segment of *première* is absent already at the lexical level, the /e/ is in a closed syllable at this level, since the word *première* is not yet followed by *ami* at this level, and the /r/ consequently cannot form a syllable with this word at the lexical level. The generalization about the surface quality of /e/ is therefore without exceptions if one takes syllable structure at the lexical level into account.

The assumption that there is lexical and a post-lexical level originates from Lexical Phonology (Kiparsky 1982; Mohanan 1986). Following van Oostendorp (1995), Booij (1997), Kiparsky (1997), and others, we incorporate this assumption in Optimality Theory, and reject the claim that Optimality Theory should assume only one level.[1]

---

[1] The assumption of a lexical level is not crucial in the present study, since the analyses that will be presented can be easily translated into analyses that assume only one level in combination with Output-Output-constraints, which require the forms of a morpheme to be identical in different words (e.g. Benua 1995; Kenstowicz 1996), or Sympathy constraints, which require the output candidate to be identical to a certain rejected candidate in some respect (McCarthy 1998). We assume a lexical level because the data which will be discussed in the present study can be most easily accounted for under this assumption.

2.3.4 Prosodic Phonology

The most important representational theory adopted in the present study is Prosodic Phonology as developed by Selkirk (1980, 1986), Nespor & Vogel (1982, 1986), and others. This theory assumes that an utterance has a hierarchical prosodic structure. We adopt the assumption that segments form syllables (σ), syllables form feet (F), feet form prosodic words (PW), prosodic words form phonological phrases, phonological phrases form intonational phrases, and intonational phrases form phonological utterances. The prosodic structure of an utterance influences its realization, since several phonological constraints and phonetic processes are sensitive to prosodic constituency.

Grammatical words which form single prosodic constituents with grammatical words on their right may be grouped in syntactic constituents with grammatical words on their left, and vice versa. This appears, for instance, form the Dutch sentence (2). The article *het* in this sentence belongs prosodically to the preceding word *kocht*, as it forms a syllable with the [t] of this word. Syntactically, however, it belongs to the following word *boek*, with which it forms an NP (or DP).

(2) An instance of non-isomorphy between (i) morphosyntactic on the one hand and (ii) syllabic or (iii) prosodic word structure on the other hand (see also Booij 1996: 219).

	Jan	*kocht*	*het*	*boek*
(i)	[$_{CP}$[$_{NP}$ jɑn]	[$_{V}$ kɔxt]	[$_{VP}$[$_{NP}$ ət	buk]]]
(ii)	(jɑn)$_{σ}$	(kɔx)$_{σ}$(t	ət)$_{σ}$	(buk)$_{σ}$
(iii)	(jɑn)$_{PW}$	(kɔxt	ət)$_{PW}$	(buk)$_{PW}$
	Jan	bought	the	book

The assignment of prosodic structure follows the same principles in all languages. The formation of syllables and feet is generally sensitive to the quality of the segments, while the formation of Phonological and Intonational Phrases and Phonological Utterances is influenced by the morphosyntactic structure of the utterance. In at least some languages, the formation of Intonational Phrases and Phonological Utterances is also influenced by speech rate, and the positions of the accents in the utterance (e.g. Gussenhoven 1988; Marsi et al. 1997).

In the unmarked case, every prosodic constituent belongs to exactly one prosodic constituent of the category immediately above it, and consists exclusively of constituents of the category immediately below it. This generalization is called the Strict Layer Hypothesis, and is enforced by constraints on domination. These constraints are violable, like all phonological constraints (Selkirk 1995: 441). As a consequence, in some languages syllables forming affixes or unstressed function words are not part of a foot, but are directly incorporated into a higher prosodic constituent (see e.g. Vogel 1994; Booij 1995; Peperkamp 1997), and segments may be incorporated directly into the prosodic word (for an example see §3.6).

Syllables, feet, and prosodic words must be present at the lexical level of phonology, since several phonological constraints which are relevant to the lexical level refer to these constituents (e.g. Booij 1988; Nespor 1990; Zsiga 1992). This is evident, for instance, from the French data discussed in section 2.3.3. The constituents that are higher in the hierarchy than the prosodic word, however, are not present before the post-lexical level. They may correspond to several grammatical words, and therefore cannot be formed before the grammatical words are strung together, i.e. before the post-lexical level.

One of the merits of Prosodic Phonology is that it can adequately account for the prosodic behaviour of clitics. Clitics are function words, like /ət/ 'it' in (2), that do not prosodically behave like normal words. They are unstressed, and often form a syllable with segments of preceding or following grammatical words. Their exceptional prosodic behaviour can be accounted for under the assumption that they do not form prosodic words of their own, but are incorporated into the preceding or following prosodic word, form a new prosodic word together with the preceding or following prosodic word, or are directly incorporated into a Phonological Phrase (Selkirk 1995).

Many phenomena can be accounted for if we assume a certain prosodic structure, and the existence of constraints which are sensitive to this structure. Naturally, a prosodic structure should only be assumed if its presence can be argued for on independent grounds.

2.4 Phonetics

2.4.1 Differences between phonology and phonetics

The output of phonology is the input to phonetics (see Figure 2.1 in §2.1). The view on the role of phonetics has changed enormously during the last decades. Early generativists defined this component as the collection of universal mechanisms which determine the realization of speech, and are the automatic results of speech physiology (Chomsky & Halle 1968: 293; Kenstowicz & Kisseberth 1979). The nasalization of vowels before nasals, the fronting of /k/ before [i], and the shortening of vowels before voiceless obstruents were assumed to be examples of phonetic processes. Since phonetic processes were seen as the automatic results of speech physiology, they were assumed not to belong to the grammar proper.

This early view of phonetics proved to be untenable (Keating 1985, 1990a; Kingston & Diehl 1994). The exact realization of a segment was found to differ per language, and in some cases to be related to other properties of that language. The latter is evident, for instance, from the fact that the amount of vowel to vowel coarticulation is probably influenced by the number of vowels functioning as phonemes in the language (Manuel 1990). Besides, no part of a realization proved to be the automatic result of speech physiology. Some realizations are preferred by speech physiology, but speakers are not physically constrained to choose them. All details of the realizations of segments are apparently represented cognitively, and under explicit control of the speaker. They are part of the grammar. If the phonetic component contained only universal and automatic mechanisms, it would be empty.

Linguists have therefore proposed a new division of labour between phonology and phonetics. Many of them, e.g. Keating (1988ab, 1990b), Pierrehumbert (1990), Cohn (1993), Jun (1995), and Zsiga (1997), assume that the difference between the phonological and phonetic component is that the former deals with symbols, while the latter relates these symbols to actual speech.

The phonological symbols, which include segments and features, represent categorical abstract, stable, timeless sounds. Phonological constraints, which refer to these symbols, consequently have categorical effects, and do not necessarily have articulatory or perceptual causes. Moreover, their relevance can be independent of speech style.

Phonetic processes are assumed to translate phonological representations into articulatory and perceptual targets, and therefore have articulatory or perceptual grounds. The articulators move from target to target. The transitions between targets, i.e. the interpolations, produce sounds whose qualities gradually change. The speed of the quality alterations depends on the type of speech: they are probably faster in careful speech, since speakers tend to realize segments more individually in more formal speech styles. Phonetic processes used to be regarded as the result of rules, which were called phonetic implementation rules. Several linguists nowadays prefer to see them as the effects of phonetic constraint interactions (see e.g. Flemming 1997).

The present study adopts this view of the distinction between phonology and phonetics, and therefore assumes that phonology and phonetics differ as summarized in Table 2.2. We do not need to choose between phonetic rules and constraints, since we will not attempt to formulate phonetic processes explicitly.

consequence, in some languages syllables forming affixes or unstressed function words are not part of a foot, but are directly incorporated into a higher prosodic constituent (see e.g. Vogel 1994; Booij 1995; Peperkamp 1997), and segments may be incorporated directly into the prosodic word (for an example see §3.6).

Syllables, feet, and prosodic words must be present at the lexical level of phonology, since several phonological constraints which are relevant to the lexical level refer to these constituents (e.g. Booij 1988; Nespor 1990; Zsiga 1992). This is evident, for instance, from the French data discussed in section 2.3.3. The constituents that are higher in the hierarchy than the prosodic word, however, are not present before the post-lexical level. They may correspond to several grammatical words, and therefore cannot be formed before the grammatical words are strung together, i.e. before the post-lexical level.

One of the merits of Prosodic Phonology is that it can adequately account for the prosodic behaviour of clitics. Clitics are function words, like /ət/ 'it' in (2), that do not prosodically behave like normal words. They are unstressed, and often form a syllable with segments of preceding or following grammatical words. Their exceptional prosodic behaviour can be accounted for under the assumption that they do not form prosodic words of their own, but are incorporated into the preceding or following prosodic word, form a new prosodic word together with the preceding or following prosodic word, or are directly incorporated into a Phonological Phrase (Selkirk 1995).

Many phenomena can be accounted for if we assume a certain prosodic structure, and the existence of constraints which are sensitive to this structure. Naturally, a prosodic structure should only be assumed if its presence can be argued for on independent grounds.

2.4 Phonetics

2.4.1 Differences between phonology and phonetics

The output of phonology is the input to phonetics (see Figure 2.1 in §2.1). The view on the role of phonetics has changed enormously during the last decades. Early generativists defined this component as the collection of universal mechanisms which determine the realization of speech, and are the automatic results of speech physiology (Chomsky & Halle 1968: 293; Kenstowicz & Kisseberth 1979). The nasalization of vowels before nasals, the fronting of /k/ before [i], and the shortening of vowels before voiceless obstruents were assumed to be examples of phonetic processes. Since phonetic processes were seen as the automatic results of speech physiology, they were assumed not to belong to the grammar proper.

This early view of phonetics proved to be untenable (Keating 1985, 1990a; Kingston & Diehl 1994). The exact realization of a segment was found to differ per language, and in some cases to be related to other properties of that language. The latter is evident, for instance, from the fact that the amount of vowel to vowel coarticulation is probably influenced by the number of vowels functioning as phonemes in the language (Manuel 1990). Besides, no part of a realization proved to be the automatic result of speech physiology. Some realizations are preferred by speech physiology, but speakers are not physically constrained to choose them. All details of the realizations of segments are apparently represented cognitively, and under explicit control of the speaker. They are part of the grammar. If the phonetic component contained only universal and automatic mechanisms, it would be empty.

Linguists have therefore proposed a new division of labour between phonology and phonetics. Many of them, e.g. Keating (1988ab, 1990b), Pierrehumbert (1990), Cohn (1993), Jun (1995), and Zsiga (1997), assume that the difference between the phonological and phonetic component is that the former deals with symbols, while the latter relates these symbols to actual speech.

The phonological symbols, which include segments and features, represent categorical abstract, stable, timeless sounds. Phonological constraints, which refer to these symbols, consequently have categorical effects, and do not necessarily have articulatory or perceptual causes. Moreover, their relevance can be independent of speech style.

Phonetic processes are assumed to translate phonological representations into articulatory and perceptual targets, and therefore have articulatory or perceptual grounds. The articulators move from target to target. The transitions between targets, i.e. the interpolations, produce sounds whose qualities gradually change. The speed of the quality alterations depends on the type of speech: they are probably faster in careful speech, since speakers tend to realize segments more individually in more formal speech styles. Phonetic processes used to be regarded as the result of rules, which were called phonetic implementation rules. Several linguists nowadays prefer to see them as the effects of phonetic constraint interactions (see e.g. Flemming 1997).

The present study adopts this view of the distinction between phonology and phonetics, and therefore assumes that phonology and phonetics differ as summarized in Table 2.2. We do not need to choose between phonetic rules and constraints, since we will not attempt to formulate phonetic processes explicitly.

To summarize, phonology and phonetics are assumed to differ as indicated in Table 2.2.

Table 2.2 Main differences between phonology and phonetics.

Phonology	Phonetics
• Symbolic representations;	• The symbolic representations of phonology are related to actual speech;
• Constraints have categorical effects;	• Mechanisms may have gradient effects;
• Constraints do not necessarily have articulatory or perceptual motivations;	• Mechanisms have articulatory or perceptual motivations;
• The relevance of a constraint can be independent of speech style.	• The effects of a mechanism are influenced by speech style.

2.4.2 Reduction of articulatory effort

When people speak less carefully, they tend to realize their speech with less articulatory effort (§1.2). They reduce the size of the articulatory gestures, and change their temporal organization. The effects of the reduction of the size of articulatory gestures cannot generally be part of the phonological form, because the phonological form is assumed to represent segments and features, not gestures. Gestures with different sizes correspond to different features only if they are contrastive in some languages, and most size differences introduced by the reduction of articulatory effort are not. The effects of the temporal reorganization of gestures in general cannot be incorporated into the phonological form either, since phonological symbols are related to abstract time units. In conclusion, the effects of gesture reduction and temporal reorganization generally have to be introduced by phonetics, and will be referred to as "phonetic". They are contrasted with effects which are already present in the phonological form, and which will be called "phonological".

The reduction of articulatory effort tends to affect segments which are not acoustically salient, or not phonologically contrastive in the relevant position in the language involved. The fact that the affected features tend to be acoustically non-salient has to be accounted for by phonetics, since this component deals with the physical aspects of sounds. The fact that they tend to be non-contrastive is a matter of phonology, as it is this component that indicates which features are contrastive.

The present study assumes that the reason why particularly the realization of non-contrastive features is affected by the speaker's natural tendency to reduce articulatory effort is that non-contrastive features, which are unspecified in the lexicon (cf. Steriade 1987; Archangeli 1988; Mester & Itô 1989; Cho 1990; Lombardi 1995a, and other work), can remain unspecified in phonology and in the input to phonetics (Keating 1990a; Cohn 1993). When a feature remains unspecified in a certain language, it is not translated into articulatory or perceptual targets, and speakers do not have to make any effort to realize the targets corresponding to one of its values. Only those targets are reached that need no additional articulatory effort in the given circumstances. Consequently, the relevant segments sometimes sound as if they are specified with the plus value, and sometimes as if they are specified with the minus value of the feature. Hence, a segment for which [voice] is not contrastive may sometimes sound as voiced in certain languages and sometimes as voiceless, depending on factors such as speech rate, and the type of adjacent segments (Chapter 7). The details of its realization in these languages are determined by ease of articulation.

2.4.3 Competing views

In addition to the view that the phonological component deals with abstract symbols, while the phonetic component deals with actual articulation and perception, there are at least two other views of the relation between phonology and phonetics. First, there is Articulatory Phonology (Browman & Goldstein 1989, 1990ab, 1992a, 1995, etc.). This theory assumes that phonemes are characterized in terms of coordinated articulatory gestures in phonology, and therefore that the difference between phonology and phonetics is much smaller than in the approach adopted in the present study. Articulatory Phonology is not adopted here, since it is entirely based on gestures, and is therefore only insightful if all the characteristics of the gestures made during the relevant realizations are exactly known, which will not be the case here.

Second, several linguists have suggested that phonological constraints can incorporate principles of articulation, aerodynamics, and perception (see for instance Ohala 1990ab; Boersma 1998; and Hale & Reiss forthcoming for arguments against this view). They argue that many types of data can only be explained adequately if constraints expressing these principles can interact with purely phonological constraints. In this view contrastive features tend to be realized faithfully because the faithfulness constraints on these features dominate the constraints which state that articulatory effort is minimal. Non-contrastive features are not always realized faithfully because the faithfulness constraints on them are sometimes dominated by the constraints on articulatory effort. The role of phonetics in this view is not worked out in as much detail.

The present study does not adopt this view of phonology, since there is not as yet a generally accepted theory incorporating it. Moreover, this view is only insightful if all the details of all relevant realizations are exactly known, as it is only then that it is possible to state which articulatory, aerodynamic and perceptual constraints dominate which faithfulness constraints. This will not be the case in the present study.

Most analyses which adopt the assumption that the phonological component deals with abstract symbols, that the phonetic component relates these symbols to actual speech, and that features can be unspecified in phonetics can be well translated into analyses adopting Articulatory Phonology or theories which allow phonological constraints to refer to articulatory, aerodynamic and perceptual principles. When they are translated into analyses within Articulatory Phonology, the unspecified segments should be regarded as not corresponding to any gesture. When they are translated into analyses adopting physical principles, the unspecified segments should be regarded as specified segments which have to satisfy faithfulness constraints that are lower in the constraint hierarchy than certain constraints expressing physical principles.

3 Phonology of Dutch

3.1 Introduction

This chapter will describe the phonological characteristics of Standard Dutch which are relevant to this study. First, the vowel and consonant inventories of this language will be discussed (§3.2 and §3.3), and then the consonant realizations which are unfaithful to the underlying representations with respect to [voice] (§3.4), or with respect to length (§3.5), and analyses of these realizations. The chapter ends with a section on prosodic constituency (§3.6).

It needs to be emphasized that the notion of Standard Dutch is not well-defined. It is generally taken to be the set of varieties which are accepted by most speakers (cf. §1.1), and is therefore quite a subjective notion.

3.2 Inventory of vowels

Dutch has sixteen native vowels: schwa, three narrow diphthongs, nine monophthongs, and three wide diphthongs. The schwa is defined here as located in the centre of the vowel space, or as having inherited the articulatory properties of the adjacent segments. This definition is in line with Brownman & Goldstein's (1992b) claim that schwa has its own [place]-specifications, which are positioned in the centre of the vowel space, and that the targets implied by these specifications are reached only if the vowel is sufficiently long.

The three narrow diphthongs are [ei], [ou], and [øy] (see e.g. Zwaardemaker & Eijkman 1928: 125, 129). Phonologically, they belong to the monophthongs. For instance, like the monophthongs, they are often reduced to schwa, whereas this is rarely the case with the wide diphthongs (Stroop 1974: 320; Kager 1989: 300; Booij 1995: 131, 134). In accordance with all previous literature, we assume that [ei] corresponds to underlying /e/, [ou] to /o/, and [øy] to /ø/. We will refer to the group of these three vowels and the actual monophthongs as the phonological monophthongs.

Table 3.1 shows all native phonological monophthongs. It characterizes them by their height, backness, and roundness. Height is characterized as a combination of the features [high] and [mid], since the vowels present four degrees of vowel height (cf. Booij 1995: 17).

Table 3.1 The native phonological monophthongs.

	[−back]		[+back]	
	[−round]	[+round]	[−round]	[+round]
[+high, −mid]	i	y		u
[+high, +mid]	ɪ, e	ʏ, ø		o
[−high, +mid]	ɛ			ɔ
[−high, −mid]			a, ɑ	

The [ʏ] occurs for instance in the words *bus* [bʏs] 'bus', and *kunst* [kʏnst] 'art' (e.g. Booij 1995; Rietveld & van Heuven 1997; Gussenhoven 1999). It is represented as /θ/ by van Reenen & Elias (1998).

Table 3.2 lists the three wide diphthongs, and their backness and roundness. The height of these vowels changes from [−high] to [+high, −mid] in the course of time.

Table 3.2. The wide diphthongs.

	backness	roundness
/ɛi/	[−back]	[−round]
/œy/	[−back]	[+round]
/ɑu/	[+back]	[+round]

The native vowels /ɪ, ɛ, ɔ, ʏ, ɑ/ and /i, y, u, e, ø, o, a, ɛi, œy, ɑu/ behave as two different groups in phonology. The groups differ, for instance, in the number of coda consonants that they can precede (see e.g. Booij 1995: 69 ff, and §3.6)).

Zonneveld (1978), Trommelen (1983), van der Hulst (1984), Booij (1995) and others argue that the differences between the two groups result from the number of time-slots to which the vowels are linked in phonology. These linguists claim that the vowels /ɪ, ɛ, ɔ, ʏ, ɑ/ are each linked to one slot, whereas the vowels /i, y, u, e, ø, o, a, ɛi, œy, ɑu/ are linked to two slots. The former vowels are therefore believed to be phonologically short, and the latter phonologically long. The assumption of phonologically long and short vowels in Dutch is supported by experiments which show that there is a relation between the phonological length of a vowel and its acoustic length: phonologically short vowels are generally acoustically shorter than phonologically long vowels realized under the same circumstances (Nooteboom 1972).

The experiments also show that the relation between phonological and acoustic vowel length does not hold for /i, u, y/, since these phonologically long vowels pattern acoustically with the phonologically short vowels in all contexts, except before /r/ (Nooteboom 1972). Moreover, the relation between phonological and acoustic vowel length appears to hold for the other vowels only when they are in the head of feet (Rietveld et al. 1999). Because of these and other facts, van Oostendorp (1995), Rietveld et al. (1999), and Gussenhoven (1999) reject the claim that /i, y, u, e, ø, o, a/ are always phonologically long. They argue that the phonological differences between /ɪ, ɛ, ɔ, ʏ, ɑ/ and /i, y, u, e, ø, o, a/ are due to the specifications of these vowels with respect to the feature [lax]: the former vowels are specified as [+lax], whereas the latter are [−lax]. The difference in acoustic vowel length between the [+lax] vowels and /e, ø, o, a/ in the head of feet would be completely due to phonetic processes (van Oostendorp 1995: 34), or to the phonological specification of /e, ø, o, a/ as long when they function as the head of feet (Gussenhoven 1999; Rietveld et al. 1999).

This study adopts the assumption that /ɪ, ɛ, ɔ, ʏ, ɑ/ and /i, y, u, e, ø, o, a/ differ in laxness. Analyses based on this assumption seem to incur less severe problems than analyses assuming differences in phonological vowel length. Whereas the most important problem for analyses adopting [lax] is the physical correlate of this feature, analyses assuming differences in phonological vowel length have to explain, among other things, why Dutch, unlike the great majority of languages, does not allow CV syllables, and treats syllables ending in long vowels as light, but closed syllables with short vowels as heavy (see van Oostendorp 1995: 28 ff.). The choice between the two views is not crucial for the present study.

In addition to the sixteen native vowels discussed so far, Dutch has a number of marginal vowels, which only occur in loan words. These vowels are long /i, y, u, ɛ, œ, ɔ, ɑ/ and nasalized /ɛ, œ, ɔ, ɑ/ (Booij 1995: 6).

3.3 Inventory of consonants

Table 3.3 lists the native consonants of Dutch.

Table 3.3 The native consonants.

	Bilabial	Labiodental	Alveolar	Palatal	Velar	Glottal
Stops	p, b		t, d		k	
Fricatives		f, v	s, z		x, ɣ	h
Nasals	m		n		ŋ	
Liquids			l, r			
Glides		ʋ		j		

Two remarks have to be made with respect to these consonants. First, there is much geographical variation in the realization of several phonemes, among which /r/ and /x/. Since the present study is not principally concerned with geographical influence, it will not distinguish among the different geographical variants. Hence, these variants are not indicated in the table.

Second, the voiced/voiceless distinction is less relevant for fricatives than for stops, as is evident from, for instance, the realization of these obstruents in word-initial position. Many speakers never realize labiodental or velar fricatives as voiced in this position, and speakers from the western part of the Netherlands, from Friesland, and from the areas near the big rivers, do not tend to realize word-initial alveolar fricatives as voiced either. They realize words that originally started with voiced fricatives in some dialects of Middle Dutch with voiceless fricatives (Collins & Mees 1981: 159; Gussenhoven & Bremmer 1983: 57; Slis & van Heugten 1989; van Reenen 1994). In contrast, the voiced/voiceless distinction for word-initial stops is respected by all speakers of Dutch.

Further support for the weakness of the opposition between voiced and voiceless fricatives is provided by the number of word pairs which only differ in the [voice]-specification of their initial fricative. There are only 9 minimal word-pairs with initial /f/ and /v/, and 10 pairs with initial /s/ and /z/ (see 1 and 2). Since some of them contain archaic or infrequent members, such as *feil* and *fat*, the actual number of minimal pairs in every-day speech is even smaller.

(1) Minimal word-pairs with initial /f/ vs. /v/.

faal	[fa:l]	'fail'	-	*vaal*	[va:l]	'faded'
fat	[fɑt]	'dandy'	-	*vat*	[vɑt]	'hold'
fee	[fe:]	'fairy'	-	*vee*	[ve:]	'cattle'
feil	[fɛil]	'failing'	-	*vijl*	[vɛil]	'file'
fel	[fɛl]	'fierce'	-	*vel*	[vɛl]	'skin'
fin	[fɪn]	'Finn'	-	*vin*	[vɪn]	'fin'
fier	[fi:r]	'proud'	-	*vier*	[vi:r]	'four'
fout	[fɑut]	'mistake'	-	*vouwt*	[vɑut]	'folds' (verb form)
fries	[fris]	'Frisian'	-	*vries*	[vris]	'freeze'

(2) Minimal word-pairs with initial /s/ vs. /z/.

saai	[sa:j]	'boring'	-	*zaai*	[za:j]	'sowing'
sagen	[sa:xən]	'legends'	-	*zagen*	[za:xən]	'to saw'
C	[se:]	'letter C'	-	*zee*	[ze:]	'sea'
sein	[sɛin]	'signal'	-	*zijn*	[zɛin]	'to be' or 'his'
sijs	[sɛis]	'siskin'	-	*zeis*	[zɛis]	'scythe'
cent	[sɛnt]	'cent'	-	*zend*	[zɛnt]	'send'
set	[sɛt]	'set'	-	*zet*	[zɛt]	'move'
Sien	[sin]	'Sien' (name)	-	*zien*	[zin]	'to see'
sier	[si:r]	'show'	-	*zier*	[zi:r]	'the least bit'
sonde	[sɔndə]	'probe'	-	*zonde*	[zɔndə]	'sin'
sul	[sʏl]	'softy'	-	*zul*	[zʏl]	'will' 2nd ps. sg.

In contrast, the number of word pairs which only differ in the [voice]-specification of their initial stop is much larger. Some pairs are listed in (3) and (4). Many pairs, such as *pak* - *bak*, and *perk* [pɛrk] - *berk*, consist of members which are highly frequent, and can occur in nearly identical semantic contexts.

(3) Some minimal word-pairs with initial /p/ vs. /b/.

paard	[pa:rt]	'horse'	-	*baard*	[ba:rt]	'beard'
pak	[pɑk]	'parcel'	-	*bak*	[bɑk]	'bin'
pad	[pɑt]	'path'	-	*bad*	[bɑt]	'bath'
part	[pɑrt]	'part'	-	*Bart*	[bɑrt]	'Bart' (name)

pauw	[pɑu]	'peacock'	-	*bouw*	[bɑu]	'building'
peer	[pe:r]	'pear'	-	*beer*	[be:r]	'bear'
perk	[pɛrk]	'bed in garden'	-	*berk*	[bɛrk]	'birch'
pek	[pɛk]	'pitch'	-	*bek*	[bɛk]	'mouth'
pont	[pɔnt]	'ferry-boat'	-	*bont*	[bɔnt]	'fur'
poot	[po:t]	'paw'	-	*boot*	[bo:t]	'boat'
pot	[pɔt]	'jar'	-	*bot*	[bɔt]	'bone'
preken	[pre:kən]	'to preach'	-	*breken*	[bre:kən]	'to break'
prul	[prʏl]	'trash'	-	*brul*	[brʏl]	'roar'
pui	[pœy]	'facade'	-	*bui*	[bœy]	'mood'

(4) Some minimal word-pairs with initial /t/ vs. /d/.

tak	[tɑk]	'branch'	-	*dak*	[dɑk]	'roof'
tas	[tɑs]	'bag'	-	*das*	[dɑs]	'scarf'
teken	[te:kən]	'sign'	-	*deken*	[de:kən]	'blanket'
teren	[te:rən]	'to live on'	-	*deren*	[de:rən]	'to harm'
tik	[tɪk]	'tap'	-	*dik*	[dɪk]	'thick'
toen	[tun]	'then'	-	*doen*	[dun]	'to do'
tol	[tɔl]	'top'	-	*dol*	[dɔl]	'crazy'
tolk	[tɔlk]	'interpreter'	-	*dolk*	[dɔlk]	'dagger'
tor	[tɔr]	'beetle'	-	*dor*	[dɔr]	'barren'
top	[tɔp]	'top'	-	*dop*	[dɔp]	'shell'
tooi	[to:j]	'decoration'	-	*dooi*	[do:j]	'thaw'
turen	[ty:rən]	'to peer'	-	*duren*	[dy:rən]	'to last'
tuin	[tœyn]	'garden'	-	*duin*	[dœyn]	'dune'
touw	[tɑu]	'rope'	-	*dauw*	[dɑu]	'dew'

The voiced/voiceless distinction is the weakest for the velar fricative, which is nearly always realized as voiceless in the western part of the Netherlands. Since the language variant spoken in this region will be the subject of the present investigations (§5.2.4.2), this study will not recognize the voiced velar fricative as a phoneme (cf. Gussenhoven 1992: 45). It will assume that velar fricatives which were underlyingly voiced in Middle Dutch are nowadays underlyingly voiceless.

The realization of the regular past-tense suffix poses a problem for the assumption that Western Dutch has no underlying /ɣ/. The phonological shape of this suffix depends on the underlying [voice]-specification of the preceding obstruent, and is different after segments which are realized as [x] and [ɣ] in some variants of Dutch (§3.4.6). We assume that the past-tense forms which are irregular under the assumption that all velar fricatives are underlyingly voiceless, such as *zaagde* [zaɣdə] 'sew' and *droogde* [droɣdə] 'dried', are a relict of earlier variants of Dutch, and that it is indicated in the lexicon that they are realized with [də].

Dutch consonants are not always realized faithfully to their underlying representations. In certain contexts they are systematically realized unfaithfully to their underlying [voice]-specifications, or their lexical lengths. This is the topic of the next two sections.

3.4 Voiced and voiceless obstruents

3.4.1 Introduction

Obstruents are not always realized faithfully to their underlying [voice]-specifications. The realizations of some are determined by the qualities of the adjacent segments, instead of their underlying specifications. That is,

- coda and word-final obstruents are realized as voiced or voiceless depending on the quality of, at least, the following segment;
- fricatives and the initial /d/s of some function words are, obligatorily or optionally, realized as voiceless after obstruents;
- the initial stop of the regular past-tense suffix is realized as voiceless after underlyingly voiceless obstruents, and as voiced in all other contexts.

Sections 3.4.2 to 3.4.6 will discuss the realizations of these obstruents, and present previous analyses. The sections form a basis especially for Chapter 7, which will propose a new analysis for the voiced/voiceless realization of obstruents.

3.4.2 Obstruents in coda positions

All obstruents in coda positions generally surface as voiced before voiced stops, and as voiceless in all other contexts. This is illustrated in (5). Example (5a) shows that the infinitive forms *verwijden* [vɛrʋɛidən] and *verwijten* [vɛrʋɛitən], with the infinitive marker *-en* (-[ən]), differ in their meanings. The feature [voice] is therefore distinctive for stem-final obstruents, and the stem of *verwijden* must be underlyingly /vɛrʋɛid/, and the stem of *verwijten* /vɛrʋɛit/. Example (5b) shows that when the underlying /d/ and /t/ of /vɛrʋɛid/ and /vɛrʋɛit/ are in coda position and followed by a voiced stop, they both correspond to [d] in the phonetic form. When they are in coda position and not followed by a voiced stop, they correspond to [t] (see examples 5cd).

(5)	a.	*verwijden*	/vɛrʋɛid-ən/	[(vɛr)$_{\sigma}$(ʋɛi)$_{\sigma}$(dən)$_{\sigma}$]	'widen-inf.'
		verwijten	/vɛrʋɛit-ən/	[(vɛr)$_{\sigma}$(ʋɛi)$_{\sigma}$(tən)$_{\sigma}$]	'reproach-inf'
	b.	*verwijdbaar*	/vɛrʋɛid-bar/	[(vɛr)$_{\sigma}$(ʋɛid)$_{\sigma}$(ba:r)$_{\sigma}$]	'widen-able',
		verwijtbaar	/vɛrʋɛit-bar/		'reproach-able'
	c.	*ik verwijd*	/ɪk vɛrʋɛid/	[(ɪk)$_{\sigma}$(vɛr)$_{\sigma}$(ʋɛit)$_{\sigma}$]	'I widen',
		ik verwijt	/ɪk vɛrʋɛit/		'I reproach'
	d.	*verwijd niets*	/vɛrʋɛid nit/	[(vɛr)$_{\sigma}$(ʋɛit)$_{\sigma}$(nit)$_{\sigma}$]	'widen not',
		ver*wijt niets*	/vɛrʋɛit nit/		'reproach not'

Coda obstruents surface as voiced before voiced stops also if their voiced variants do not constitute phonemes in Dutch. This appears from examples as (6). Example (6a) shows that the verb stem of *maken* ends in a /k/. Example (6b) shows that this /k/ is realized as [g] when it is in coda position before a voiced stop, even though the /g/ is not a phoneme of Dutch.

(6)	a.	*maken*	/mak-ən/	[(ma:)$_{\sigma}$(kən)$_{\sigma}$]	'make-inf.'
	b.	*maak dit*	/mak dɪt/	[(ma:g)$_{\sigma}$(dɪt)$_{\sigma}$]	'make this'

Many analyses have been proposed for the realization of word-final obstruents in coda positions as voiced or voiceless. The analyses formulated within SPE-rules assume that all word-final obstruents (Trommelen & Zonneveld 1979: 60 ff.; Berendsen 1983, 1986: 46; Zonneveld 1983: 298 ff.) or all obstruents which are in coda position at the end of the lexical level (Booij 1981: 42, 79) are devoiced by a rule called Final Devoicing. Those obstruents which are followed by voiced stops are subsequently voiced by a rule of Regressive Voice Assimilation. Final Devoicing and Regressive Voice Assimilation were formulated by Booij (1981: 42, 79) as in (7) and (8). The symbol "$" indicates a syllable-boundary.

(7) Final Devoicing

$$[-\text{son}] \rightarrow [-\text{voice}] \; / \; \_\_ \; \$$$

(8) Regressive Voice Assimilation

$$[-\text{son}] \rightarrow [+\text{voice}] \;/\; \_\_\ \$ \begin{bmatrix} -\text{son} \\ -\text{cont} \\ +\text{voice} \end{bmatrix}$$

Analyses formulated within Optimality Theory were proposed by Lombardi (1995b, 1999), Grijzenhout & Krämer (1999), and Mascaró & Wetzels (1999). Lombardi's (1995b, 1999) analysis has been the most influential. It starts from the assumption, defended in Lombardi (1995a), that [voice] is a privative feature, i.e. that only voiced obstruents are specified for [voice] in phonology. It adopts the constraints formulated in (9).

(9) IDENTONSET(LARYNGEAL) (abbreviated IDONSLAR):

Consonants before sonorants should be faithful to underlying laryngeal specification.

IDENT(LARYNGEAL) (abbreviated IDLAR):

Consonants should be faithful to underlying laryngeal specification.

*LAR:

Consonants should not have laryngeal features.

AGREE:

Obstruent clusters should agree in voicing.

Constraint ranking (10) ensures that coda obstruents are realized as voiced before voiced stops, and as voiceless in all other contexts.

(10) IDONSLAR, AGREE >> *LAR >> IDLAR

This is illustrated in Tableaux 3.1 and 3.2.

Tableau 3.1 Obstruent in utterance-final position.

bed /bɛd/ 'bed'	IDONSLAR	AGREE	*LAR	IDLAR
(bɛd)$_\sigma$			**!	
☞ (bɛt)$_\sigma$			*	*
(pɛt)$_\sigma$	*!			**

The output candidate [pɛt] in Tableau 3.1 is non-optimal for the input /bɛd/ because being [p]-initial, it violates the undominated constraint IDONSLAR, whereas the [b]-initial candidates [bɛd] and [bɛt] do not. These latter candidates violate *Lar or IDLAR. The form [bɛd] violates *LAR twice, as it contains two voiced obstruents. The candidate [bɛt] violates *LAR only once. Since *LAR dominates IDLAR, [bɛt] is optimal.

Tableau 3.2 Obstruent before a voiced stop.

kijkdag /kɛik-dɑx/ 'view day'	IDONSLAR	AGREE	*LAR	IDLAR
(kɛik)$_\sigma$(dɑx)$_\sigma$		*!	*	
☞ (kɛig)$_\sigma$(dɑx)$_\sigma$			**	*
(kɛik)$_\sigma$(tɑx)$_\sigma$	*!			*

The input /kɛik-dɑx/ in Tableau 3.2 has as optimal output candidate [kɛigdɑx]. This candidate violates *LAR twice, because the coda obstruent of the first syllable and the onset obstruent of the second syllable are both linked to the laryngeal feature [+voice]. The violations are felicitous, since they avoid violations of AGREE and IDONSLAR. AGREE is satisfied because the obstruents are identically specified for [voice]. The constraint IDONSLAR is satisfied because the onset obstruent is realized in accordance with its underlying [voice]- specification.

The analyses mentioned so far — the ones formulated within SPE-rules as well as the ones formulated within Optimality Theory — assume that all obstruents which are realized as voiced are specified as [+voice] in the phonological form, and that all obstruents which are realized as voiceless are either specified as [-voice] or unspecified for [voice], depending on whether they adopt underspecification theory. In addition, these analyses assume that final obstruents are always voiced before voiced stops, and voiceless in all other contexts.

A different type of analysis is proposed by Slis (1985). Following the classical generative analyses, Slis assumes that coda obstruents undergo Final Devoicing, and are phonologically specified as [-voice]. Unlike the classical generative analyses, his analysis does not assume a phonological process of regressive voice assimilation. Slis claims that coda obstruents are not always realized as voiced before voiced stops, and if they are, this is the result of phonetic coarticulation.

Neither Slis' analysis nor the analyses formulated within Optimality Theory deal with the voiced realization of some word-final obstruents before certain vowel-initial function words. These voiced realizations are the topic of section 3.4.3.

3.4.3 Word-final obstruents before vowels

The final obstruents of at least some words can be realized as voiced and voiceless before certain vowel-initial function words. The final /b/ of *heb* /hɛb/ 'have', for instance, can be realized as [b] and [p] before *ik* /ɪk/ 'I' ([hɛpɪk], [hɛbɪk]). The analyses discussed in section 3.4.2, with the possible exception of the one proposed by Slis, do not predict the voiced realizations. They are designed to have voiced word-final obstruents in their outputs only before voiced obstruents.

Berendsen (1986), Booij (1985, 1987, 1995), and Gussenhoven (1986) have proposed additional analyses for the realization of word-final obstruents as voiced or voiceless before vowels. These analyses are all formulated in terms of phonological rules, and assume Final Devoicing. They differ in their outputs, as is illustrated in Table 3.4. This table shows the realization of post-vocalic word-final /d/s and /t/s according to the three analyses. The table is restricted to stops which precede the function words *ik* /ɪk/ 'I', *het* /ət/ 'it', *er* /ər/ 'there', and *ie* /i/ 'he', and do not occur in word-combinations stored in the lexicon. It holds exclusively for Standard Dutch.

Table 3.4 The realization of word-final coronal stops, according to Berendsen, Booij, and Gussenhoven. The stops are divided into categories characterized by their underlying [voice]-specification and the type of following function word.

Category	Example		Realization according to		
			Berendsen	Booij	Gussenhoven
/...Vt̲ ɪk/	*weet ik*	'know I'	[t]	[t]	[t] and [d]
/...Vd̲ ɪk/	*had ik*	'had I'	[t] and [d]	[t]	[t] and [d]
/...Vt̲ ət/	*weet het*	'know(s) it'	[t] and [d]	[t]	[t]
/...Vd̲ ət/	*had hat*	'had it'	[t] and [d]	[t]	[t]
/...Vt̲ ər/	*weet er*	'know(s) there'	[t]	[t]	[t]
/...Vd̲ ər/	*had er*	'had there'	[t]	[t]	[t]
/...Vt̲ i/	*weet ie*	'knows he'	[t]	[t]	[t]
/...Vd̲ i/	*had ie*	'had he'	[t] and [d]	[t]	[t]

Berendsen (1986) claims that word-final obstruents which are underlyingly voiced can surface as voiced and voiceless before every vowel-initial clitic except *er* /ər/ 'there'. Underlyingly voiceless obstruents, on the other hand, normally surface as voiceless before such clitics. They can surface as voiced only if they are /t/s and the clitic is schwa-initial.

Berendsen accounts for his data as follows. He assumes that clitics have to be incorporated into the preceding prosodic word, or directly adjoined to the preceding or following phonological phrase. When clitics are incorporated into the preceding prosodic word, they belong to the same syllabification domain as the preceding word-final obstruents. Vowel-initial clitics then form syllables with these preceding word-final obstruents, which means that these obstruents end up in onset position, and cannot be devoiced by Final Devoicing. They surface as voiced if they are underlyingly voiced, and as voiceless if they are underlyingly voiceless. When clitics are directly adjoined to the preceding or following phonological phrase, they do not form syllables with the preceding obstruents. These obstruents consequently remain in coda position, and are devoiced by Final Devoicing. They surface as voiceless, independently of their underlying [voice]-specifications. In other words, underlyingly voiced obstruents before vowel-initial clitics surface as voiced or voiceless, depending on the type of prosodic incorporation, but underlyingly voiceless obstruents always surface as voiceless. There is one exception to this generalization: when schwa-initial clitics are directly adjoined to the preceding or following phonological phrase, preceding /t/s can be voiced by a special phonological rule called "clitic /t/-voicing".

Booij (1985, 1987, 1995) claims that all word-final obstruents are normally realized as voiceless in Standard Dutch, although there are a few host+clitic combinations in which underlyingly voiced word-final obstruents can be realized as voiced.

Booij's account rests on two assumptions: first, that Final Devoicing applies at the end of the lexical level, and second, that lexical phonology generally deals with hosts and clitics in isolation. These assumptions imply that the final obstruents of nearly all grammatical words are in coda position at the lexical level, and are devoiced by Final Devoicing. They surface as voiceless, independently of their underlying [voice]-specifications, and independently of the prosodic status of the following grammatical word. However, some special host+clitic combinations are present as units at the lexical level (Booij 1987: 222). If the host of such a combination ends in an obstruent, and the clitic starts with a vowel, the host-final obstruent forms a syllable with the clitic at the lexical level. This obstruent is consequently in onset position at the lexical level, and is not devoiced by Final Devoicing. It is realized faithfully to its underlying [voice]-specifications. Host+clitic combinations which behave as units at the lexical level are those stored in the lexicon.

One such host+clitic combination is the sequence *heb ik* /hɛb ɪk/ 'have I'. It is probably stored as it is highly frequent, and is sometimes realized as [hɛk]. The contraction [hɛk] cannot result from productive synchronic phonological processes of Dutch, since this type of reduction can only apply to a restricted number of host+clitic combinations (see also §9.5.2). This means that [hɛk], and therefore possibly also *heb ik* as a unit, is stored in the lexicon.

Finally, Gussenhoven (1986: 186, 187) claims that in Standard Dutch all post-vocalic obstruents can be voiced before *ik* /ɪk/ 'I', and that post-vocalic fricatives can also be voiced before other vowel-initial function words. His explanation for these realizations rests on the assumption that all word-final obstruents undergo Final Devoicing, and that the post-vocalic ones can subsequently be voiced by Intervocalic Voice Assimilation when they precede a prosodic word boundary and certain vowel-initial function words. Which types of obstruents can undergo Intervocalic Voice Assimilation before which function words is dialect-specific. There are, however, no dialects in which Intervocalic Voice Assimilation can voice obstruents preceding the form *ie* /i/ 'he'. This form is obligatory enclitic, and is therefore always incorporated into the preceding prosodic word, which means that it is never separated from the preceding obstruent by a prosodic word-boundary. This preceding obstruent consequently never satisfies the conditions for Intervocalic Voice Assimilation.

Evaluating Berendsen's, Booij's, and Gussenhoven's analyses is impossible without valid data on the realization of word-final obstruents before vowel-initial function words. Such data will be provided in Chapter 9, and they will appear to be in accordance with none of the analyses. An alternative analysis will be presented in Chapter 7.

3.4.4 Fricatives in onset positions

Fricatives in onset positions form another type of obstruent which is not always realized faithfully to its underlying [voice]-specification, as both underlyingly voiced and voiceless fricatives are realized as voiceless after obstruents. This is illustrated in (11). The indefinite article *een* can be followed by voiced as well as voiceless fricatives. Since the words *zee* and *vat* both surface with voiced fricatives after *een*, they apparently have voiced fricatives in their underlying representations. These underlyingly voiced fricatives are realized as voiceless in *diepzee* and *handvat*, in which they follow obstruents.

(11) a.	*een zee*	/ən ze/	[ənze:]	'a sea'
	een C	/ən se/	[ənse:]	'an C'
	diepzee	/dipze/	[dipse:]	'deep sea'
b.	*een vat*	/ən vɑt/	[ənvɑt]	'a grip'
	een fat	/ən fɑt/	[ənfɑt]	'a dandy'
	handvat	/ɦɑndvɑt/	[ɦɑntfɑt]	'hand grip'

Classical generative analyses (see e.g. Trommelen & Zonneveld 1979: 60 ff.; Booij 1981: 42, 79; Berendsen 1983, 1986: 46; Zonneveld 1983: 298 ff.) account for the voiceless realizations of fricatives after obstruents with a rule of Progressive Voice Assimilation. This rule makes fricatives conform to the [voice]-specifications of the preceding obstruents, which are voiceless as a result of the rule of Final Devoicing (cf. §3.4.2). Booij's (1981) formulation of the rule of Progressive Voice Assimilation can be found in (12). The symbol "$" indicates a syllable-boundary.

(12) Progressive Voice Assimilation

$$\begin{bmatrix} -\text{son} \\ +\text{cont} \end{bmatrix} \rightarrow [-\text{voice}] \,/\, \begin{bmatrix} -\text{son} \\ -\text{voice} \end{bmatrix} \; \$ \; \_$$

Within Optimality Theory, two types of analyses have been proposed for the onset fricatives. Lombardi (1995b, 1999) ascribes the voiceless realizations of these fricatives after obstruents to the high ranking of a constraint requiring fricatives to be voiceless after obstruents. Grijzenhout & Krämer (1999) and Mascaró & Wetzels (1999) ascribe the voiceless realizations to the high rankings of a constraint requiring obstruents in a cluster to agree in [voice], and a constraint requiring coda obstruents to be voiceless. Grijzenhout & Krämer (1999) and Mascaró & Wetzels (1999) assign different names to these constraints, and We will refer to them as "CONFORM" and "FINDEV". The high ranking of CONFORM has the effect that an obstruent and a following fricative are both realized as either voiced or voiceless. Because of FINDEV, they are realized as voiceless (see Tableau 3.3).

Tableau 3.3 Onset fricative after an obstruent.

handvat /hɑnd-vɑt/ 'hand grip'	CONFORM	FINDEV
(hɑnd)$_\sigma$(vɑt)$_\sigma$		*!
(hɑnt)$_\sigma$(vɑt)$_\sigma$	*!	
☞ (hɑnt)$_\sigma$(fɑt)$_\sigma$		

Since all these analyses can explain the data, choosing between them seems to be a matter of personal preference.

3.4.5 Word-initial /d/s

In contrast with initial fricatives, initial stops are generally realized faithfully to their underlying representations, also after obstruents. There is one exception: the initial /d/ of a number of function words.

When following obstruents, the /d/ of these words is, more or less, optionally realized as [t]. This is illustrated in (13). The lexical word *dien* and the function words *daar* and *dat* are obligatorily realized with [d] in utterance-initial position (13a). When they follow obstruents, the lexical word *dien* is always realized with [d] (13b), whereas the function words *daar* and *dat* are sometimes realized with [d] and sometimes with [t] (13c). The preceding obstruents are realized as voiceless when the /d/ is realized as voiceless, and as voiced when the /d/ is realized as voiced.

(13)	a.	*Dien op*	/din ɔp/	[dinɔp]	'Dish up'
		Daar	/dar/	[da:r]	'There'
		Dat	/dɑt/	[dɑt]	'That'
	b.	*opdienen*	/ɔp dinən/	[ɔbdinən]	'to dish up'
	c.	*heb daar*	/hɛb dar/	[hɛpta:r], [hɛbda:r]	'have there'
		maak dat	/mak dɑt/	[ma:ktɑt], [ma:gdɑt]	'fix that'

Zonneveld (1982, 1983) is the only linguist who has proposed analyses for the realization of word-initial /d/. His first analysis (Zonneveld 1982) assumes that /d/-initial words which always surface with [d] only have lexical forms with /d/, whereas words which can surface with [t] after obstruents have one lexical form with /d/, and one with /ð/. The forms with /d/ always surface with [d], and preceding obstruents surface as voiced as the result of Regressive Voice Assimilation (which was described in §3.4.2). The forms with /ð/ surface with a voiceless alveolar after obstruents because of Final Devoicing (described in §3.4.2) and Progressive Voice Assimilation (described in §3.4.4), and with a voiced alveolar in all other contexts. This alveolar /ð/ surfaces as a stop, as the result of a strengthening rule. This is illustrated in (14).

(14)	Underlying form:	/pd/	/pð/	/bd/	/bð/	/nð/
	Final Devoicing:			pd	pð	
	Regressive Voice Assimilation:	bd		bd		
	Progressive Voice Assimilation:		pθ		pθ	
	Strengthening:		pt		pt	nd
	Phonological and phonetic form:	[bd]	[pt]	[bd]	[pt]	[nd]

Zonneveld's second analysis (Zonneveld 1983) is very different. This analysis claims that a /d/-initial word which can be realized with [t] after obstruents is optionally incorporated into the preceding prosodic word. When it is incorporated, and follows an obstruent, the initial /d/ is part of a word-medial obstruent cluster. Word-medial obstruent clusters are voiceless by default, and therefore the /d/ is realized as [t]. In contrast, when words are not incorporated into the preceding prosodic word, their initial /d/ is realized as [d].

Zonneveld (1983) represents an improvement on Zonneveld (1982) in that it does not assume that segments which always surface as [d]s or [t]s are underlyingly fricatives. It is a step back in that it implies that the realizations of underlyingly voiced fricatives and function word-initial /d/s are due to different processes, although the two types of obstruents are both realized as voiceless after obstruents and as voiced in all other contexts. The only difference between the two types is that fricatives are obligatorily realized as voiceless, whereas the voiceless realization of word-initial /d/s seems to be more or less optional.

3.4.6 The initial stop of the regular past-tense suffix

The final type of obstruent which is realized as voiced in some contexts and as voiceless in others is the initial stop of the regular past-tense suffix, which consists of a coronal stop and a schwa. The stop is realized as voiceless after obstruents which are underlyingly voiceless, and as voiced after all other phonemes. This is illustrated in (15a-e), which show the underlying forms of certain verb stems, and the realizations of the corresponding past-tense forms. Only after /x/ the realization of the suffix is unpredictable (see 5fg).

(15) a.	*open*	/opən/	'open'	-	*opende*	[opəndə]	'opened'
b.	*zwaai*	/zʋaj/	'wave'	-	*zwaaide*	[zʋa:jdə]	'waved'
c.	*krab*	/krɑb/	'scratch'	-	*krabde*	[krɑbdə]	'scratched'
d.	*raap*	/rap/	'pick up'	-	*raapte*	[ra:ptə]	'picked up'
e.	*bak*	/bɑk/	'bake'	-	*bakte*	[bɑktə]	'baked'
f.	*lach*	/lɑx/	'laugh'	-	*lachte*	[lɑxtə]	'laughed'
g.	*zaag*	/zax/	'saw'	-	*zaagde*	[za:ɣdə]	'sew'

The realization of the suffix has been accounted for by three different analyses. The first analysis was proposed by Trommelen & Zonneveld (1979) and Zonneveld (1982). It assumes that the stop of the suffix is underlyingly /ð/. When this fricative follows a voiceless obstruent, it is devoiced by Progressive Voice Assimilation, like any other fricative. When it follows another type of phoneme, it remains voiced. The fricative is turned into a stop before phonetic implementation. This analysis produces correct outputs only if Final Devoicing does not devoice the underlyingly voiced obstruents which precede the past-tense morpheme. According to Trommelen & Zonneveld (1979) and Zonneveld (1982), this condition is met, since Final Devoicing only affects word-final obstruents (§3.4.2).

The second type of analysis has been proposed by Wetzels (1982:125), Booij (1995: 62,) and Grijzenhout (1999). It assumes that the coronal stop of the past-tense suffix is underlyingly unspecified for [voice]. Wetzels (1982) assumes that the stop always borrows, and Booij (1995) that it is always linked to the [voice]-specification of the preceding segment. Grijzenhout (1999) assumes that it is linked to the [voice]-specification of the preceding segment only if this segment is an obstruent. If the preceding segment is a sonorant, the stop is assigned the default [voice]-specification, which is [+voice] in this context.

The most recent analysis is the one proposed by Borowski (forthcoming). Borowski assumes that the initial stop of the regular past-tense morpheme is /d/. This /d/ is realized as [t] after underlyingly voiceless obstruents, because the obstruents in a cluster have to agree for [voice], and it is worse for segments of a base than for segments of an affix to be unfaithful to their underlying [voice]-specifications.

All three analyses explain the fact that some velar fricatives are followed by [də], and some by [tə], with the assumption that those followed by [də] are underlyingly voiced, and that those followed by [tə] are underlyingly voiceless. As mentioned in 3.3, we will not consider the voiced velar fricative to be a phoneme of Dutch. Instead, the past-tense forms ending in the velar fricative followed by [də] will be assumed to be stored in the lexicon. Hence, we will assume that it is indicated in the lexicon that the past-tense forms of verbs such as *zagen* /zax-ən/ 'saw-inf.' and *drogen* /drox-ən/ 'dry-inf.' are realized with [də] instead of [tə].

3.5 Geminates

Obstruents are frequently realized unfaithfully to their underlying representations not only with respect to their [voice]-specifications, but also with respect to their durations. That is, most consonants with a lexical length of two segments ("geminates") are generally realized with a duration that is shorter than the acoustic duration of two separate segments (cf. Martens & Quené 1994; Booij 1995: 68). Geminates of which the parts belong to the same prosodic word are obligatorily realized with the acoustic length of a single segment (see 16a), and geminates of which the parts are separated only by a prosodic word boundary, as in morphological compounds, are often realized with an acoustic length that is nearly as short as the length of a single consonant (see 16b). Only geminates of which the parts are separated by a prosodic word as well as a phonological phrase boundary, as for instance in (16c), are frequently relatively long.

(16) a.	*lasster*	/lɑs+stər/	[lɑstər]	'weld+er (fem.)'
b.	*kerkklok*	/kɛrk+klɔk/	[kɛrklɔk], [kɛrk·lɔk], ([kɛrk:lɔk])	'church bell'
c.	*at tonijn*	/ɑt tonɛin/	[ɑt:onɛin], [ɑt·onɛin], ([ɑtonɛin])	'ate tuna'

The relative acoustic length of geminates of which the parts are separated by at least a prosodic word boundary is influenced by speech rate. The geminates are relatively shorter when speech rate is higher.

These generalizations on length do not only hold for genuine geminates, but also for clusters of obstruents which differ only in their underlying [voice]-specifications (see e.g. 17). Such clusters generally sound as completely voiced when they end in an underlyingly voiced stop, and as voiceless in all other cases.

(17) a.	*klapband*	/klɑp+bɑnd/	[klɑbɑnt], [klɑb·ɑnt], ([klɑb:ɑnt])	'flat tyre'
b.	*waszak*	/ʋɑs+zɑk/	[ʋɑsɑk], [ʋɑs·ɑk], ([ʋɑs:ɑk])	'laundry bag'
c.	*had toch*	/hɑd tɔx/	[hɑt:ɔx], [hɑt·ɔx], ([hɑtɔx])	'had yet'

Classical generative analyses explain the relative acoustic lengths of consonants by means of a phonological rule called "degemination". This rule deletes one of two identical consonants in a sequence. It applies obligatorily within prosodic words, and optionally across word-boundaries (Booij 1995: 68). Degemination is ordered at least after Final Devoicing (which is described in §3.4.2), Regressive Voice Assimilation (described in §3.4.2), and Progressive Voice Assimilation (described in §3.4.4). The interaction of these four rules produces the correct output forms for genuine geminates, and for clusters of obstruents differing in their underlying [voice]-specification (see 18).

(18)				
Underlying form:	/bp/	/pb/	/sz/	/zs/
Final Devoicing:	pp			ss
Regressive Voice Assimilation:		bb		
Progressive Voice Assimilation:			ss	
Degemination:	p	b	s	s
Phonological and phonetic form:	[p]	[b]	[s]	[s]

Since these classical generative analyses explain the realizations of geminates by a phonological, and therefore categorical (§2.4.1), rule which deletes one consonant, i.e. one time-slot, they assume that underlying geminates always have acoustic lengths of either one or two single segments. This is contrary to fact for geminates of which the parts do not belong to the same prosodic word (see above). The classical generative analyses do therefore not adequately account for the data.

3.6 Prosodic structure

Prosodic constituency (§2.3.4) is nowadays considered an important characteristic of phonological representations. This study will refer to some characteristics of the syllable, the foot, and the prosodic word.

The relevant assumptions with respect to the prosodic syllable are that it consists of an onset, a nucleus, and a coda, and that in Dutch these constituents have the characteristics listed in (19).

(19) Relevant characteristics of the Dutch syllable (see e.g. van der Hulst 1984; Booij 1995)

- The onset contains as many consonants as possible (The Maximal Onset Principle). The maximum is 3.
- The nucleus contains a vowel.
- The coda contains maximally 1 consonant if the nucleus is a diphthong or a tense vowel. It contains 1 or 2 consonants if the nucleus is a lax vowel.

Because lax vowels have to be followed by at least one consonant, and onsets should contain as many consonants as possible, pre-vocalic obstruents following lax vowels are ambisyllabic: they belong to the preceding syllable as well as to the following one (van der Hulst 1985). For instance, the [k] in the word *akker* [ɑkər] 'field' forms, simultaneously, a coda for the syllable headed by [ɑ] and an onset for the syllable headed by [ə].

Most syllables are part of feet. The most important assumption for this study with respect to feet is that they generally consist of one or two syllables, and of three syllables in only a small number of monomorphemic words, or when several unstressed syllables which are part of different morphemes happen to be adjacent within the same prosodic word. When feet consist of more than one syllable, the left-most syllable is the head, and is the only one which bears stress (Gussenhoven 1993, and references cited there).

A prosodic word in Dutch may contain footed and non-footed syllables, and an appendix. It minimally consists of one foot, which forms its head. All coronal obstruents at the end of a prosodic word are directly incorporated into the prosodic word, and are part of neither syllables nor feet. They are said to form the "appendix" (Booij 1995: 26 ff).[1] The appendix explains why the number of consonants that can follow the last vowel of a prosodic word exceeds the number of coda consonants allowed in non-final syllables. Figure 3.1 illustrates the structure of prosodic words in Dutch with the word *herfst* [hɛrfst] 'autumn'. Since prosodic words start with a syllable, or a foot, which in turn starts with a syllable, its left-hand boundary is obligatorily aligned with a syllable-boundary, and, vice versa, the left-hand side of the first syllable is obligatorily aligned with the left-hand boundary of the prosodic word.

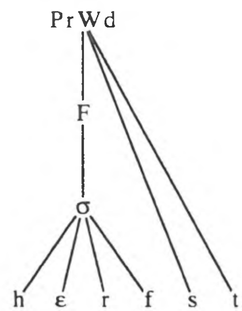


Figure 3.1 The prosodic structure of [hɛrfst].

Prosodic words may be smaller than grammatical words. First, the constituents of morphological compounds generally form prosodic words of their own (Booij 1995: 47 ff.). The compound *brand-oefening* /brɑnd-ufənɪŋ/ (['brɑnt,ufənɪŋ]) 'fire practice', for instance, consists of the prosodic words *brand* and *oefening*, as is evident from the realization of the /d/ as [t], which indicates that the /d/ is in coda position, and is therefore followed by a syllable-boundary. The syllable-boundary is expected if the two constituents of the compound belong to different prosodic words. The /d/ is then followed by a prosodic word boundary, and consequently by a syllable-boundary. If the constituents of the compound belong to the same prosodic word, the syllable-boundary would be unexpected. The two constituents then belong to the same syllabification domain, and the /d/ would be expected to form the onset

[1] This type of appendix should not be confused with appendices which contain unfooted syllables, and were proposed by Gussenhoven (1993).

of the syllable headed by the /u/, because of the Maximal Onset Principle.

In addition, the prosodic structure of *brandoefening* follows from the syllables which bear stress. These syllables also bear stress when the two constituents of the compound are realized in isolation (['(brɑnt)$_σ$], ['(u)$_σ$(fə)$_σ$(nɪŋ)$_σ$]). This suggests that the constituents form prosodic domains of their own.

Finally, the prosodic structure of *brandoefening* follows from the position of the primary stress. This stress falls on (brɑnt)$_σ$, and is therefore at a distance of three syllables from the right edge of the word. This is highly exceptional for words which are not morphological compounds (see e.g. van der Hulst 1984: 235).

The second type of constituent which does not form a prosodic word with the remainder of the grammatical word is instantiated by some exceptional suffixes such as *-achtig* '-ish' and *-baar* '-able' (Booij 1995: 30). The words containing these affixes behave like prosodic compounds with respect to e.g. syllabification. This is clear from an example like *groenachtig* /xrun-ɑxtɪx/ 'green-ish', which is realized as [(xrun)$_σ$(ʔɑx)$_σ$(tɪx)$_σ$]. The inserted glottal stop indicates that the /ɑ/ is not preceded by an onset, as glottal stops are never inserted in the middle of syllables in Dutch. Hence, the [ɑ] is preceded by a syllable-boundary in the phonological form. This boundary prevents the syllabification from satisfying the Maximal Onset Principle, and therefore must be due to the presence of a prosodic word boundary.

Finally, prefixes do not form prosodic words with their bases. They form syllabification domains of their own (Booij 1995: 30). This is evident from an example like *ontaard* /ɔnt-ard/ 'degenerate', consisting of the prefix /ɔnt-/ 'de-' and the stem /ard/ 'nature'. This word is realized with a glottal stop before *aard* ([(ɔnt)$_σ$(ʔart)$_σ$]), which implies that the [t] of *ont* does not belong to the same syllable as the [a] of *aard*. The syllable-boundary separating the two phonemes violates the Maximal Onset Principle, and therefore must be due to the presence of a prosodic word boundary. Hence, the prefix *ont* and the stem *aard* do not belong to the same prosodic word.

Prosodic words cannot only be smaller, but also larger than grammatical words. They can contain a content word plus one or more unstressed function words (Booij 1995: 170 f.f.). Examples are *weet ie* /ʋet-i/ 'knows he', and *koop het* /kop ət/ 'by it', which form single syllabification domains. The incorporated unstressed words are called clitics (§2.3.4).

For other characteristics of prosodic constituency in Dutch, which are not relevant to this study, see Booij (1995) and Marsi et al. (1997), and references cited there.

Part III

Type of data

4 Intuitions and speech as linguistic evidence

4.1 Introduction

Given the background information presented in Part II, we can now proceed with the actual study of this book. First, we will determine which type of data is most valuable given the research questions formulated in section 1.2. Then, in Parts IV and V, we will actually try to answer these questions on the basis of the data.

Data for phonological and phonetic investigations are typically provided by poetry, word games, production mistakes, sound changes, speakers' intuitions, and actual speech (cf. Ohala 1986). The first four data types can only be of secondary importance to the studies on the relevance of the speaker's tendency to reduce articulatory effort on casual speech, and therefore the present study.

The problem with poetry and word games is that one never knows what it is exactly that is responsible for the observed patterns: whether it is the underlying forms of the words, their phonetic forms, or their orthographies. They therefore rarely provide information on which properties of realizations are phonological, and which ones are phonetic in nature.

Production mistakes are realizations which do not result from the regular phonological and phonetic processes. They are unsuitable as the main data for the present study because they are difficult to recognize: they are seldom corrected in casual speech, and cannot be recognized by their characteristics, as it is unknown which characteristics are ungrammatical.

Finally, sound changes probably do provide valuable data, but there are too few of them to form a sufficient base to study casual speech. Sound changes, therefore, cannot form the main data of the present study, either.

It seems, then, that the main data for this study must be provided by speakers' intuitions or samples of actual speech. As speakers' intuitions form the basis of many linguistic studies of all types, they may also be of use when studying realizations in casual speech. Actual speech will certainly provide valuable data, since it can show the regular realizations of segments and words, and indirectly provide information on which realizations are due to phonology or phonetics (see e.g. Ohala 1981).

This chapter will discuss speakers' intuitions (§4.2) and actual speech (§4.3 and §4.4) as possible data types in more detail. We will conclude that the data for the investigations of the present study should come from a corpus of spontaneous conversations (§4.5). Chapter 5 will introduce the corpus that was developed for this study.

4.2 Linguistic intuitions

4.2.1 General aspects

Many linguistic analyses are based on speakers' intuitions. One of the reasons is that intuitions are generally easily available: they can be obtained without having to leave one's arm-chair. Theoretical studies are particularly likely to be based on intuitions also because most theoretical linguists consider competence rather than performance to be the focus of their research (see e.g. de Saussure 1916: 37; Chomsky 1964: 26), and assume that intuitions reflect the competence of speakers better than any other data type.

Intuitions as linguistic evidence have enormously increased our knowledge of the grammars of languages (Newmeyer 1983: 49). The main reason for this is that intuitions, in contrast to all other types of data, can provide positive and negative evidence: speakers generally know which realizations are grammatical and which ones are not.

Linguistic intuitions are not accepted as valuable data by all linguists. Some critics wonder whether intuitions indeed reveal characteristics of the linguistic competence, or only provide data for a theory of linguistic intuitions (Levelt 1972: 22). Other critics argue that intuitions may well be influenced by the way they are elicited, and by the speakers' expectations (Labov 1975: 26).

Finally, there is the idea that, whereas speakers may have clear intuitions about the lexical forms of words, they are rather ignorant of actual realizations (Mohanan 1986: 194). This idea implies that studies of realizations in casual speech, to name an example, should not be based mainly on intuitions.

Clear evidence that speakers do not have valid, i.e. correct, intuitions on the non-lexical characteristics of words is provided by the realization of word-final obstruents before vowel-initial function words in Dutch. This will be shown in section 4.2.2.

4.2.2. Intuitions on word-final obstruents

The realization of word-final obstruents as voiced or voiceless depends on the quality of the initial segment of the following word (§3.4.1 to §3.4.3)), and is therefore a non-lexical property. It could therefore provide an excellent testing ground for the idea that speakers' intuitions about non-lexical properties are not valid.

We saw in section 3.4.3 that various models have been proposed to deal with the realization of word-final obstruents as voiced or voiceless before vowel-initial function words. These analyses differ from each other particularly in their outputs. Table 3.4, here repeated as Table 4.1, shows the various generalizations that these outputs allow to be made about the realization of post-vocalic coronal stops before the function words *ik* /ɪk/ 'I', *het* /ət/ 'it', *er* /ər/ 'there', and *ie* /i/ 'he' in Standard Dutch.

Table 4.1 The realization of word-final coronal stops, according to Berendsen, Booij, and Gussenhoven. The stops are divided into categories characterized by their underlying [voice]-specification and the type of following function word.

Category	Example		Realization according to Berendsen	Booij	Gussenhoven
/...Vt̲ ɪk/	*weet ik*	'know I'	[t]	[t]	[t] and [d]
/...Vd̲ ɪk/	*had ik*	'had I'	[t] and [d]	[t]	[t] and [d]
/...Vt̲ ət/	*weet het*	'know(s) it'	[t] and [d]	[t]	[t]
/...Vd̲ ət/	*had het*	'had it'	[t] and [d]	[t]	[t]
/...Vt̲ ər/	*weet er*	'know(s) there'	[t]	[t]	[t]
/...Vd̲ ər/	*had er*	'had there'	[t]	[t]	[t]
/...Vt̲ i/	*weet ie*	'knows he'	[t]	[t]	[t]
/...Vd̲ i/	*had ie*	'had he'	[t] and [d]	[t]	[t]

It is clear from Table 4.1 that Berendsen, Booij and Gussenhoven have conflicting views about the possible realizations of these stops. As their data have been gathered mostly from their own intuitions, supplemented with incidental observations, it seems that their intuitions differ significantly, which supports the idea that speakers' intuitions about the realization of these obstruents may not be valid.

There are, however, several additional possible reasons why linguists' intuitions may differ. The three linguists might be speaking, and hence describing, different geographical varieties. Booij comes from the eastern part of the Netherlands, whereas Gussenhoven comes from the western part. Moreover, the researchers may be subconsciously influenced to some extent by the predictions of their own linguistic models.

To test the hypothesis that speakers' intuitions on the realization of word-final obstruents before vowels , and therefore on the non-lexical properties of realizations, are invalid, we conducted two small experiments in which we elicited the intuitions of a number of linguistically "naive" speakers of the same variety of Dutch who had few or no theoretical expectations.

First, we asked ten first-year students of linguistics about their intuitions on the stops indicated in Table 4.1. These students were all born and raised in the western part of the Netherlands, and were unaware of any analysis of the realization of word-final obstruents before vowel-initial function words. They were female with ages ranging from 19 to 47. We asked them orally, and all at the same time. On the basis of their answers, it was possible to distinguish two separate groups. One group claimed that all word-final stops indicated in Table 4.1 are sometimes realized as voiced and sometimes as voiceless. The other group thought that these stops are always realized as voiceless. These groups, then, appeared to have different intuitions, but it is not easy to see why, as group membership did not correlate with factors such as age or geographical background, and the students had no preconceived ideas about linguistic analyses of word-final obstruents before vowels. The students' answers suggest that speakers' intuitions on the realization of word-final obstruents are invalid.

Interestingly, one student expressed doubts about the assumption that word-final obstruents can be realized as voiced before vowel-initial function words by saying *Ik geloof niet dat ik dat doe* 'I don't believe I do that', while realizing *dat ik* /dɑt ɪk/ 'that I' as [dɑdɪk], as was in fact noticed and remarked on by all the other students. This clearly shows that at least some speakers do not have access to explicit knowledge about their own linguistic behaviour. In conclusion, this small experiment provides evidence for the hypothesis that intuitions on word-final obstruents, i.e. on non-lexical properties of realizations, are not valid.

We then tested the hypothesis that speakers' intuitions on the non-lexical properties are invalid in a more formal way by interviewing sixteen male subjects with academic degrees who had lived all their lives in the western part of the Netherlands (Subjects A to P). These subjects' dates of birth, past and present places of residence, and professions can be found in Appendix A. They were selected particularly because they spoke approximately the same variety of Dutch (§5.4.3). Moreover, they were unaware of any analysis of the realizations of stops in intervocalic positions, and could therefore not be expected to be strongly influenced by any preconceived ideas. We asked these sixteen subjects to fill in questionnaires about their intuitions on the realizations of the word-final and word-medial stops of some words and word-combinations. Appendix B lists their answers.

Since the realizations of the word-medial stops are independent of the preceding or following word, they can be, and probably are, determined at the lexical level. Hence, if speakers have valid and consistent intuitions on the realizations of lexical properties, and invalid and inconsistent intuitions on non-lexical properties, they will have identical intuitions on the realizations of the word-medial stops, and different intuitions on the realizations of the word-final stops.

With respect to the word-medial obstruents, the sixteen subjects appeared to have nearly identical intuitions: they generally assumed that these obstruents are realized in accordance with their underlying [voice]-specifications. Apparently, then, speakers have consistent intuitions on the lexical properties of words.

There was less agreement among the subjects with respect to the word-final obstruents. According to the intuitions of seven subjects, these obstruents are generally realized as voiceless (Subjects D, E, G, J, L, and N), or sometimes as voiced and sometimes as voiceless (Subject C). The intuitions of five other subjects (Subjects B, H, I, M, and O) implied that the realization of word-final stops is mainly determined by the type of the following function word. These subjects generally intuited that stops before *ik* can be realized as voiced, whereas those before *ie* are always voiceless (see Table 4.2 overleaf). Finally, the intuitions of the remaining subjects implied that the realization of word-final stops is influenced by the underlying [voice]-specifications (see Table 4.3 overleaf). Particularly Subjects A, K, and P intuited that underlying /d/s are more often realized as voiced than /t/s.

Table 4.2. The intuitions of Subjects B, H, I, M, and O on the realizations of word-final coronal stops. The stops are divided into categories according to their underlying [voice]-specifications, and the type of following function word.

Category	Possible realization according to			
	Subject B	Subject H	Subject I	Subjects M, O
/...Vt̲ ɪk/	[d]	[t] or [d]	[t] and [d]	[t] and [d]
/...Vd̲ ɪk/	[d]	[t] or [d]	[t] and [d]	[t] and [d]
/...Vt̲ ət/	[t] and [d]	[d]	[t]	[t]
/...Vd̲ ət/	[t] and [d]	[d]	[t]	[t]
/...Vt̲ ər/	[t] and [d]	[t] and [d]	[t] and [d]	[t]
/...Vd̲ ər/	[t] and [d]	[t] and [d]	[t] and [d]	[t]
/...Vt̲ i/	[t] and [d]	[t]	[t]	[t]
/...Vd̲ i/	[t]	[t]	[t]	[t]

Table 4.3. The intuitions of Subjects A, F, K, and P on the realizations of word-final coronal stops. The stops are divided into categories according to their underlying [voice]-specifications, and the type of following function word.

Category	Possible realization according to			
	Subject A	Subject F	Subject K	Subject P
/...Vt̲ ɪk/	[t] or [d]	[d]	[t]	[t] or [d]
/...Vd̲ ɪk/	[d]	[d]	[t] and [d]	[t] or [d]
/...Vt̲ ət/	[t] or [d]	[d]	[t] and [d]	[t]
/...Vd̲ ət/	[d]	[t] and [d]	[d]	[t] or [d]
/...Vt̲ ər/	[t]	[t]	[t]	[t]
/...Vd̲ ər/	[t] or [d]	[t] and [d]	[t] and [d]	[d]
/...Vt̲ i/	[t]	[t]	[t]	[t]
/...Vd̲ i/	[t] or [d]	[t] and [d]	[t] and [d]	[d]

The sixteen subjects, then, displayed nearly identical intuitions on the realization of word-medial stops, whereas they did not with respect to the realization of word-final stops. As these differences in opinion cannot solely be due to different language varieties, or different theoretical expectations, they constitute additional support for the hypothesis that intuitions are not valid with respect to those characteristics of a realization that are not encoded in the lexicon.

In summary, the results of both experiments suggest that intuitions do not constitute valid data for studies on post-lexical or phonetic properties of realizations. This does not imply that intuitions on these properties do not provide any interesting information at all, but simply that this information is of a different type. For instance, since the intuitions displayed by the subjects of the two experiments are conflicting, they support the assumption that the realization of word-final obstruents as voiced or voiceless is neither determined in the lexicon nor at the lexical level of phonology. Moreover, these intuitions suggest that the word-final obstruents are not voiced at the lexical level, since they show that many speakers believe that they are never voiced. The obstruents must therefore be voiceless at this level, or, since [-voice] is the unmarked value (see e.g. Mester & Itô 1989; Cho 1990; Lombardi 1995a), be unspecified for [voice].

4.3 Speech

4.3.1 Introduction

Section 4.2 argued that speakers' intuitions do not constitute valuable data for studies on post-lexical and phonetic processes, and therefore for studies on casual Dutch. Since collections of poetry, the results of word games, speech errors, and sound changes cannot constitute the main body of data either (§4.1), such studies can only be based on actual speech.

Analysing speech is not simple. Each analysis implies that the relevant stretches of speech have to be interpreted in several ways. First, they have to be interpreted as strings of symbols representing their perceptual characteristics. Then, it has to be determined which stretches of speech can count as actual independent data, i.e. are not the result of production mistakes, or dependent on other realizations. Finally, their characteristics have to be classified as phonological or phonetic if this distinction is important to the proposed model. Sections 4.3.2 to 4.3.5 will discuss these four types of interpretation.

4.3.2 Transcription

Stretches of speech constitute valuable data for linguistic analyses only if their perceptual characteristics are transcribed as strings of symbols, such as phonological features or characters from the International Phonetic Alphabet (International Phonetic Association 1999). The phonetic transcription of some characteristics can at least be partly based on acoustic measurements. This is, for instance, the case with tone or intonation, of which the transcription can be partly based on measurements of fundamental frequency (F_0), although F_0 does not have a one-to-one relation with perceived pitch. Naturally, transcription can be based on acoustic measurements only if it is known exactly to which acoustic properties of the sound signal the relevant perceptual characteristics are related in which context, and which are the relative strengths of these acoustic properties under different conditions. This is the case with only a few characteristics, and does not include the feature [voice] for intervocalic obstruents (§7.3.4). The upshot is that acoustic measurements have only a small part to play in the transcription of casual speech, and are of only limited use in a study like the present one.

Transcribing utterances by ear is not an easy task. The transcriber must take note of all phonetic details produced by the speaker, and decide which symbol should be used to represent which sound. It is therefore easy to make mistakes, and the task requires great concentration. Moreover, transcribing by ear is difficult because listeners normally determine which word is realized not only on the basis of what they perceive (analytic listening), but also on the basis of what they expect (semantic listening). While making phonetic transcriptions, transcribers should disregard all the expectations that automatically follow from their knowledge of the phonotactics of the language, the spelling of the word (Cucchiarini 1993: 55), its lexical representation, its pronunciation in formal speech, and so on. Ignoring one's expectations is difficult. Vieregge (1987: 9) argues that it is even impossible. According to him, phonetic transcriptions are always influenced by the transcriber's expectations, and are never objective reflections of reality.

Since transcribing by ear is so difficult, auditory transcriptions can only be held to have some validity if they have been arrived at by several independent judges. Note, however, that even if many transcribers agree on a certain transcription, this does not imply that it is valid (e.g. Cucchiarini 1993: 10). Transcribers can all be influenced by the same expectations, and therefore agree in their transcription, even though this transcription is incorrect.

One may wonder what to do with stretches of speech on which judges disagree. One possibility is to replay these stretches and see whether the judges are willing to agree on one transcription. This method probably does not yield a priori a more valid transcription, as the judges, when listening for the second time, know each other's transcriptions, and can be influenced by them, so that the transcription which is eventually accepted may not the best one, but the one obtained form the most confident transcriber.

Another possibility is to discard the problematic stretches from the data set. This method is unproblematic if the number of stretches that have to be discarded is relatively low. If the transcribers disagree about many utterances, as will often be the case for utterances realized in rapid and fluent speech (Keating 1998: 40), the method may be less adequate. The number of transcriptions that remain can be too small to warrant firm conclusions, and it is possible that discarding problematic stretches results in the removal of a more or less complete category, and consequently the loss of interesting types of data.

In summary, perceptual characteristics of stretches of speech can be transcribed on the basis of acoustic measurements if it is known exactly to which acoustic properties they are related under which conditions. Otherwise, the transcriptions have to be made by ear. Since transcribing by ear is difficult, auditory transcriptions are preferably made by several transcribers. Stretches of speech on which the transcribers do not agree should be left out of the data base if they do not constitute an important part of the data.

4.3.3 Classification as regular or irregular

Grammatical analyses should have the correct realizations as their direct outputs, which implies that for the construction and evaluation of such models, relevant stretches of speech not only have to be transcribed, but also have to be classified as to whether they result from the regular processes of the language or from production mistakes.

This classification is not a simple matter if the realizations originate from casual speech. In careful speech, both regular realizations and production mistakes can be more or less identified by their characteristics, while mistakes can also be identified by the comments of the speaker. In casual speech, this is not possible, since we do not know what the regular characteristics of realizations are in the first place, and speakers seldom correct their mistakes in this speech style. This means that the classification of realizations from casual speech can be based for the most part only on their frequency of occurrence, with the majority of cases probably constituting the regular forms, and the minority (i.e. the outliers) the mistakes. This classification method is somewhat problematic, since not all realizations with low frequencies are mistakes. Moreover, it is unknown what is the threshold frequency between realizations that are possibly mistakes and those that are probably not.

4.3.4 Classification as dependent or independent

Analyses of speech should predict the relevance of certain processes. For the construction and evaluation of such analyses, realizations therefore have to be classified as to whether they are dependent or independent. Dependent realizations of units are not computed on the basis of their separate lexical or underlying forms or those of the morphemes that they contain, but directly retrieved from the lexicon, which means that they may possess a certain characteristic not because it is the result of a phonological or phonetic process, but because it is stored in the lexicon. Hence, if the relevance of a certain process has to be determined, only independent realizations should be taken into consideration.

The classification of realizations as dependent or independent may also be problematic at times. If the assumption that realizations are temporarily stored after they have been uttered is correct, realizations in repetitions may be dependent: a speaker may retrieve the second realization of a unit in a repetition from memory, instead of computing it anew. Other examples of dependent realizations are those which are part of the same lexically stored string of words, such as tokens of the word *hand* in the expression *I am an old hand at that.* These realizations may be computed from the lexical forms of the strings in which morphemes may be stored with deviant characteristics, instead of from the underlying forms of the containing morphemes. To ensure that an investigation focuses on realizations which are computed from the underlying or lexical representations of their parts, then, any realizations in repetitions, in set phrases or phrases that are highly frequent or have unpredictable properties should be discarded. This method cannot guarantee that there are no dependent realizations in the data set, since realizations may be temporarily stored for some time and it has not been established yet that only high frequency strings may have lexical representations.

4.3.5 Classification as phonological or phonetic

For some analyses it is necessary to make a distinction between the phonological and the phonetic characteristics of a realization. Classifying characteristics as phonological or phonetic requires detailed and statistical research, instead of incidental observations.

In some cases, the classification calls for acoustic measurements. The classification of the voiceless realization of vowels is a case in point. Vowels are sometimes realized as voiceless between voiceless obstruents in several languages (see e.g. Ohala 1983; Rodgers 1998). For instance, the first vowel of the string *pick all people* is sometimes realized as voiceless in British English. These voiceless realizations may have phonetic grounds. The vocal folds generally do not vibrate during the realization of voiceless obstruents, whereas they must vibrate during the realization of vowels. This means that for the realization of a string consisting of a voiceless obstruent, a voiced vowel, and a voiceless obstruent, the vocal folds have to start and end vibrating during the realization of the vowel and the beginning of the following obstruent. This is relatively difficult if the relevant vowel is acoustically very short, and acoustically short vowels can therefore sound as voiceless between voiceless obstruents on phonetic grounds. Hence, if one wants to know the nature of the voiceless realizations of vowels between obstruents, one should determine whether only acoustically short vowels are voiceless. If this is the case, the realization may be purely phonetic. An actual classification which was based on acoustic measurements can be found in Ohala (1981).

The classifications of some other characteristics as phonological or phonetic requires articulatory data (see e.g. Cohn 1993; Zsiga 1995). The perceptual absence of [t] before [m] is a case in point. Browman & Goldstein (1989: 216) showed on the basis of articulatory data that when the [t] is inaudible in the phrase *perfect memory*, this may be due in at least some cases to the fact that it is realized during the closure of the [m] (cf. §1.1). The [t] is then articulatory present although its acoustical consequences are absent, and its absence is phonetic in nature.

4.4 Two types of speech

4.4.1 Introduction

It may be obvious that for speech data to be valid and verifiable, they have to be recorded on tape. Relevant stretches of speech can then be replayed several times, which increases the probability that they are interpreted correctly. Moreover, they can then be listened to by several linguists at different moments, which is important when they have to be transcribed by ear.

There are roughly two types of recorded speech: unscripted, as realized in spontaneous conversations, and scripted, such as sentences read aloud. We will refer to recordings of unscripted speech as "speech corpora" and to recordings of scripted speech as "recordings from experiments".

Both types of recordings have their advantages, of which the most important ones will be described in sections 4.4.2 and 4.4.3. Which type of recording is most adequate depends on the type of study.

4.4.2 Advantages of speech corpora

Speech corpora have several advantages over recordings from experiments. One advantage is that they offer the best chances of containing casual speech. Speakers who are recorded for a corpus perform a natural task: they convey a message in their own words. As a result, they generally do not take long to get used to the fact that their speech is being taped, and mostly speak as they would do under normal conditions. In contrast, speakers in experiments perform less natural tasks, since they have to read aloud sentences, often in some pre-agreed format. Because of these tasks, they are more or less constantly reminded that their speech is being tape recorded, and, as a consequence, tend to tailor their speech more in conformity with some standard, or in conformity with what they think the researcher requires. They speak casually less often.

A second advantage of corpora is that they can form an excellent basis for a survey of the possible realizations of phonemes and words, including those realizations of which little is known. Corpora generally contain a diversity of types of words and combinations of words, and therefore illustrate nearly all phenomena of connected speech. Recordings from experiments, on the other hand, generally illustrate only the process which is the subject of study. This process can never be completely unknown, since the sentences that the subjects have to read aloud have to be designed so as to contain the contexts that are maximally relevant to the process. The designer of an experiment therefore has to know at least the application domain of the process under investigation.

Finally, consulting an existing corpus is less time-consuming than designing and executing new experiments. Corpora illustrate many processes, and can therefore form the basis of surveys of very different types of realizations. A corpus compiled for a survey of the realization of /t/, for instance, can also be consulted for research on the realization of /n/, the realization of full vowels as schwas, and so on. If a good corpus is available, linguists may save themselves a lot of time and effort by consulting it, whether they intend to use it for a full-scale quantitative investigation or only for an exploratory or preliminary pilot study.

4.4.3 Advantages of recordings of experiments

Recordings of experiments have two important advantages over corpora. One advantage is that it is relatively easy to retrieve the realizations of relevant items from recordings of experiments. These recordings consist of sentences which are all constructed so as to contain the relevant items, which means that it is known beforehand which sentences contain which items. This is of course not true for corpora. As a consequence, collecting relevant realizations from recordings of experiments is less time consuming.

The most important advantage of recordings of experiments is that they can be made to contain exactly the data which are relevant to the study. The speakers are instructed precisely what to say and how to speak, and consequently generally provide exactly the required data. Speakers in corpora, on the other hand, are fairly free in their talk, and even if they have received instructions as to how to talk and what to talk about, they may in fact not realize the required words or not speak in the required way.

This has two implications. First, recordings of experiments are better suited to studies of low-frequency words or types of realization, since such words and realizations are by definition underrepresented in normal conversations. Kennedy (1998: 34) states that it takes a corpus of at least 500 million words realized during conversations with widely differing topics, that is, a much larger corpus than any currently available or planned, to achieve a reasonable number of low frequency words. In contrast, recordings from experiments with many fewer words can contain as many tokens of low frequency words and realizations as the researcher thinks are necessary.

Second, recordings of experiments are better suited to analyses which require several variables to remain constant. Analyses of the influence of a certain factor on the acoustic durations and intensities of units are a case in point. Acoustic characteristics are influenced by variables such as speech rate, intonation, and the mood of the speaker, which generally do not have the same values for all units in normal conversations. When comparing units from corpora, one therefore has to take the influence of these variables into account. This is difficult, if not impossible, because no analysis for the influence of all the variables is available. In contrast, if the acoustic characteristics of units from experiments are compared, these variables do not have to be taken into account, since they can be made to remain constant for the relevant units in the experiment.

4.4.4 Summary

The principal advantages of speech corpora and recordings from linguistic experiments are listed in Table 4.4. Speech corpora are especially useful in studies of casual speech, and in studies surveying various processes, including processes of which little is known, whereas experiments provide valuable data particularly for studies of low-frequency words and realizations and for studies which require several variables to remain constant.

Table 4.4 The principal advantages of speech corpora and recordings of experiments.

Speech corpora	Recordings of experiments

• They offer the best chances of containing casual speech;
• They are suitable for surveys of several processes, including processes of which little is known;
• The same corpus is suitable for very different types of studies.
	• Retrieval of relevant realizations is relatively easy;
	• They provide exactly the data required for a particular study and are therefore valuable in particular for studies
	  • of low-frequency words;
	  • which require variables to remain constant.

4.5 Summary and consequences for the present study

This chapter started with a short discussion of various types of data for linguistic research. It concluded that the only types which can possibly form the main data for studies of casual speech are speakers' intuitions and actual speech. The remainder of this chapter focussed on these two types.

The discussion of linguistic intuitions concluded that intuitions can form the main body of data only for studies on the lexical properties of words. As casual speech is influenced by post-lexical and phonetic processes, this implies that the present study should not be based on intuitions. It will accordingly be based on recordings of actual speech.

Stretches of speech can function as data for linguistic studies when their perceptual characteristics are transcribed as strings of symbols. Most perceptible characteristics, including the voiced/voiceless realizations of obstruents, can only be transcribed by ear. This study will therefore be based on such auditory transcriptions, and in order to minimize mistakes, most transcriptions will be made by several transcribers.

There are roughly two types of recordings of speech: corpora with unscripted speech, and recordings from linguistic experiments with scripted speech. As corpora are especially suitable for research on casual speech, and for general surveys, the present study will be based on a corpus, in the hope that this corpus will provide enough data of the right type for statistical analyses. The corpus will be described in the next chapter.

5 A corpus of casual Standard Dutch

5.1 Introduction

Chapter 4 concluded that the evidence for this study has to come from a corpus. For the purposes of the present investigation, such a corpus should meet the following requirements.

1. The corpus should contain casual speech.
2. The corpus should represent Standard Dutch, and therefore contain all varieties that are considered as standard (§1.1).
3. The subjects should speak the same variety of Standard Dutch. Statistical analyses produce more reliable results when they are based on more data. The results of the study are therefore more valid if the realizations of the different subjects can be pooled, which can only be done if the speakers do not differ with respect to their use of the realizations at issue. The risk that they do will be minimized if they speak exactly the same language variety.
4. The recordings that make up the corpus should be readily available to ensure that the realizations of its utterances can be studied in detail.
5. The corpus should be free of background noise to facilitate the transcription.
6. The corpus should be easy to work with, and therefore should not be too large.

7. The corpus should not be too small, as it has to contain enough realizations for the purposes of the present investigation. Unfortunately, the size minimally required could not be calculated in advance on the basis of the frequencies of occurrence of the relevant segments in the relevant contexts, since it was not known exactly which were the relevant contexts. The percentage of the realizations that would not be transcribed unanimously by the phoneticians consulted, and that therefore would have to be excluded for the analyses, was also an unknown property.

8. Every speaker should be represented by enough realizations to ensure that the corpus is suitable for studies of inter-speaker variability. Such studies are needed in order to assess what should be regarded as a characteristic of Standard Dutch, and what as a characteristic of the idiosyncratic language variety of a certain speaker.

The requirement that the corpus should contain only one language variety (Requirement 3) conflicts with the requirement that it should represent Standard Dutch (Requirement 2), as a corpus cannot simultaneously contain a single and several language varieties. One way out of this dilemma is to give up the requirement that the corpus should represent Standard Dutch, and to focus on just a single variety that is regarded as standard by most speakers of Dutch. This solution was not adopted here, because only people who are born in the same geographical area and are seldom in touch with people from other areas can be assumed to be speakers of the same single variety. There are only few, if any, such speakers in the Netherlands, and the corpus consequently would contain a variety spoken by only a very limited number of speakers. This would mean that the data from the corpus could not be compared to and supplemented by data on Standard Dutch from the literature, as such data represent slightly different varieties, to say the least. This would diminish the value of the investigations based on the corpus.

Another possible solution for the conflict between Requirements (2) and (3), and the one that will be chosen here, is to compromise between the two requirements, and build a corpus that contains several similar varieties of Standard Dutch. Studies of inter-speaker variability will then have to determine whether data from the different speakers can be pooled for the analysis of a certain realization. If they can, the results of analyses run on these data can be assumed to hold for a large number of speakers, and may be compared with data from the general literature on Standard Dutch.

Since compiling corpora is very time-consuming, we investigated whether there were any corpora available which met the requirements. This did not appear to be the case. The existing corpora of conversations we know of contain various, and very different, variants of Dutch (e.g. van der Wijst's 1996 corpus), contain dialects (e.g. Elias' 1977 corpus), or cannot be consulted because their tapes are no longer available (e.g. Uit den Boogaart's 1975 Eindhoven corpus).

This is why we decided to compile a new corpus for the purposes of this study. The choice for its structure and the type of speakers will be motivated in section 5.2. The recording process and the speakers on each recording will be described in section 5.3. Finally, the corpus will be evaluated in section 5.4, and section 5.5 will describe its present form.

5.2 Design

5.2.1 Introduction

A corpus is characterized at least by

1. its overall structure.
2. the place where the recordings are made.
3. the elicitation methods used.
4. the structure of the recordings.
5. the conversation topics.
6. the type of speakers.

The characteristics of our corpus will be described in sections 5.2.2 to 5.2.4. Section 5.2.2 will discuss the decisions that were taken with respect to characteristics (1-3), while section 5.2.3 will discuss the decisions with respect to the structure of the recordings and the conversation topics (characteristics 4 and 5). Section 5.2.4, finally, will deal with the selection of the speakers (characteristic 6).

5.2.2 General characteristics of the corpus

The general characteristics of the corpus were primarily determined on the basis of Requirements (1), (2), and (5-8). First, a solution had to be found for the conflict between Requirements (1) and (4). Requirement (4) states that the corpus should be free of background noise, and therefore more or less dictates that the corpus should be recorded in a soundproof room. The requirement that the corpus should contain casual speech (Requirement 1), on the other hand, more or less dictates that the corpus should not be recorded in such a room, since people in soundproof rooms are very much aware of the fact that their speech is being observed, and therefore tend to speak more formally than usual (Labov 1972: 209).

Labov (1972: 209) mentions three methods to elicit casual speech from people who know that their speech is under observation. One method entails that the speakers are encouraged to talk before and after they fulfil certain tasks. People unconsciously assume that their speech is not being observed at these moments, and consequently tend to talk more naturally. This method of eliciting casual speech is especially effective if the corpus is designed to contain only a small number of utterances. If the corpus is designed to contain a large number of utterances, like the corpus needed for this study (Requirement 7), the method is less effective, since it implies recording many hours of speech over and above the actual recordings that are made for the corpus itself. Labov's first method of eliciting casual speech must therefore be supplemented by other methods if it is to be used for the compilation of the corpus for the present study.

A second method mentioned by Labov is involving the speakers in topics which recreate strong emotions they have felt in the past. This method elicits casual speech since people do not pay much attention to the way they speak when they are in an emotional state. Which question elicits emotions depends on the speakers' cultural backgrounds and their characters. A question which makes Americans emotional is, for instance, "Have you ever been in a situation where you were in serious danger of being killed" (Labov 1972: 92). It is very difficult to come up with enough good questions to keep speakers emotional for a long time. Since the corpus needed for this study has to contain many utterances of every speaker (Requirement 8), this implies that it cannot solely consist of utterances that were realized while the speakers were in an emotional state.

Finally, a speaker is encouraged to speak casually when he talks to people he knows, instead of to an interviewer: the presence of an acquaintance will make him forget more easily that his speech is under observation, and may restrain him from showing that he is intimidated by the circumstances. Recordings of group conversations, however, are of little use to studies which require narrow phonetic transcriptions, because they generally contain much background noise and simultaneous and overlapping talk, which make them difficult to transcribe.

We decided to record the corpus in a soundproof room, and to use a combination of all of Labov's three methods for eliciting casual speech in order to overcome the problems associated with each of them. Casual speech was elicited by having a speaker talk to a friend, or a close colleague, on topics which were intended to take their minds off their speech. In addition, casual speech was elicited by encouraging the speakers to talk freely before and after the "official" conversations.

Some of the "official" conversations took place in our presence, which we do not think inhibited the subjects in any way or prevented them from speaking casually, since we all had similar social backgrounds (§5.2.4.2), and can therefore be considered to be peers. It would have been ideal if we could have found subjects with the required characteristics (§5.2.4) in our own circle of acquaintances, but this proved impossible. The strictest requirement that we could, and therefore did, meet was that at least one speaker of each pair was either an acquaintance of ours, or an acquaintance's acquaintance. If he was the acquaintance of an acquaintance, our mutual acquaintance was present at the recording, if possible.

Each pair of speakers was recorded for 90 minutes. The requirement that the corpus contains casual speech (Requirement 1) calls for lengthy recordings, because people tend to speak more naturally after they have been in a certain situation for a while. Lengthy recordings also meet the requirement that the corpus should contain a relatively large number of realizations by every speaker (Requirement 8). In contrast, the requirement that the corpus contains the speech of several people (Requirement 2) calls for short recordings, since the number of speakers in a corpus is inversely proportional to the durations of the recordings: if the recordings are shorter, the corpus can contain the speech of more people. Finally, the recordings should not be too long, as it is difficult to make people speak continually for a long time. A time-span of 90 minutes is probably a good compromise between these various requirements.

Similarly, the number of recordings represents a compromise between various requirements, namely Requirement (6), which requires the corpus to be easy to handle, and therefore to be small, Requirement (7), which requires the corpus to be large, so that it contains enough realizations for the intended investigations, and Requirement (2), which requires the corpus to represent Standard Dutch, and therefore to consist of the speech of several subjects. A number of recordings was expected to prove unsuitable for inclusion in the corpus because it was likely that some subjects would turn out not to meet the requirements formulated in section 5.2.4, or not to speak casual Standard Dutch. We estimated that if eight to ten recordings turned out to be usable, the corpus would meet all requirements as well as possible. For this reason, we decided to record 10 pairs of speakers.

Summing up, then, the corpus we compiled consists of conversations between two friends or close colleagues in a soundproof room. We made ten recordings of 90 minutes in all. Casual speech was elicited from the speakers mainly by having them discuss the topics described in section 5.2.3.

5.2.3 Structure of the recordings and conversation topics

Each recording session consisted of three parts. The first two parts were meant to elicit casual speech from the subjects, while the third was meant to elicit realizations which can be used in possible future acoustic research on vowel reduction. Casual speech was elicited in two parts, that is with two different tasks, since it seemed unlikely that a single task could make people speak continuously for almost 90 minutes. There is a further advantage to be gained from having two tasks: they give the subjects the opportunity to speak between the "official" tasks, at which moments their talk tends to be particularly casual (§5.2.2).

The first part of each recording session consisted of completely free conversations between the subjects, in the presence of the author and in the presence of our mutual acquaintance. The speakers chose their own topics, and the author, or our mutual acquaintance, sometimes asked questions or made some remarks in order to keep the conversations going. This part of the recording session lasted approximately 40 minutes.

Then the speakers were left alone for the second part of the recording session, in which they had to play a role-play. This role-play contained a part in which the speakers had to negotiate about the purchase of camping goods. One speaker played the salesman of camping equipment, while the other one played a camping store owner. The salesman had to sell tents, sleeping-bags, and back-packs for prices as high as possible. He was told that the goods come in packages of one hundred only, and that he could not sell the products separately. The store owner was instructed to purchase a maximum of 75 sleeping-bags and 75 back-packs for prices as low as possible, and not to buy tents. The salesman only knew the production costs of his products, while the store owner only knew the prices of the products in his shop. This negotiation task was inspired by van der Wijst (1996: 118).

Before the negotiation task was incorporated into the role-play, we tested it a number of times in a pilot experiment to see whether it elicits natural and casual speech. The speakers received their instructions on paper, just before the recording sessions, and these instructions were improved on after each session. The negotiations turned out to elicit speech that sounds natural and casual to us, provided the speakers are instructed that the salesman and the store owner are friends, and therefore use the more intimate forms. The negotiation task appeared to be adequate for the compilation of corpora of casual speech.

In the pilot experiment, the negotiation task took on average only 20 minutes. Since the first two parts of the recording sessions together had to fill nearly 90 minutes, the role-play had to last longer. This is why the negotiations were incorporated into a social visit. In the resulting role-play, the two speakers discussed several topics before and after the negotiations.

The speakers were free to choose their own topics of conversation during these parts of the role-play, except that they had to start with a discussion of a party which they both visited, and where the guests discussed issues such as euthanasia, the education policy, religion, and so on. If they had not attended such a party together, they had to pretend that they had. The party plot suggests many possible topics of conversations, which is why it was expected to help the speakers getting started. It also offered a range of topics which are more or less controversial, and therefore increased the probability that the speakers would become emotional, and be monitoring their speech less.

If the speakers happened to dry up, the following topics were suggested in their instructions: amusement parks, holidays, visits to family/friends, TV-programmes, pets, cinema, and the performance of their sports clubs. The salesman was encouraged to talk at length about dentists, whereas the shop owner was encouraged to talk about driving tests. Conversations about dentists and driving tests tend to arouse strong emotions, which is why the speakers were encouraged to discuss both topics. The probability that they would do so was assumed to be larger if one topic was suggested to one speaker, while the other one was suggested to the other speaker. Hence, the instructions for the salesman and the store owner slightly differed with respect to the suggested conversation topics.

People can talk freely and naturally only if they do not have to hide anything. The speakers were therefore allowed to talk about their own occupations. They were told that they could imagine being just part-time salesmen or store owners, and exercise their own profession for most of the time.

This version of the role-play was tried in a second pilot experiment. None of the speakers in this experiment had participated in the first one. The speakers received written instructions, again just before the recording sessions. These instructions resembled the ones given in Appendix C, and were improved on after every performance of the role-play. The performances showed that it is particularly the conversations that take place before the negotiations that make this version of the role-play even more acceptable for inclusion in our corpus than the first one.

These conversations extend the length of the role-play by at least 15 minutes, and cause the subjects to realize speech which sounds even more natural and casual to us. Since this version of the role-play appeared to yield good results, it was adopted for the actual recording sessions for the corpus.

After the free conversations and the role-play, the speakers were asked to realize the (non-) words *hit* [hɪt], *het* [hɛt], *hot* [hɔt], *hut* [hʏt], *hat* [hɑt], *hiet* [hit], *huut* [hyt], *hoet* [hut], *heet* [het], *heut* [høt], *hoot* [hot], and *haat* [hat] as carefully as possible (part 3 of the recording sessions). These words represent all phonological monophthongs of Dutch (§3.2), and the careful realizations of these vowels are necessary for possible future acoustic research on vowel reduction.

To sum up, each recording consists of three parts: free conversations in the author's presence, a role-play, and the realization of all Dutch vowels in monosyllabic words by every speaker. The remarks made by the speakers after each part were recorded as well. The recordings therefore have the structure indicated in Table 5.1.

Table 5.1 The structure of the recordings.

Speakers' activity	Average duration	Author present
• Free conversations (including final remarks)	40 minutes	yes
• Role-play consisting of - free conversations - negotiations - free conversations	40 minutes	no
• Spontaneous conversations	3 minutes	yes
• Realization of monosyllabic words	5 minutes	no
• Spontaneous conversations	2 minutes	yes

5.2.4 The speakers

5.2.4.1 Introduction

The validity of a corpus naturally depends on the type of speaker. In order to ensure that the corpus meets Requirements (2) and (3) as well as possible, it was decided to select subjects who speak similar varieties of Standard Dutch (§5.1). This was done by ensuring that their geographical and social backgrounds as well as their personal characteristics were similar (§5.2.4.2 and §5.2.4.3), and by incorporating their speech into the corpus only if it could be classified as Standard Dutch (§5.2.4.4).

5.2.4.2 Geographical and social backgrounds

The speakers for the corpus were selected from people who had lived all their lives in the western provinces of North-Holland, South-Holland, and Utrecht. We decided that the speakers should come from these regions because the dialects spoken in these provinces do not differ greatly (Nerbonne 1999), and therefore the varieties of Standard Dutch spoken there are probably similar as well. Moreover, people from the western provinces were obvious subjects since the corpus was compiled in Amsterdam, which is in the west of the Netherlands. Finally, speakers from North-Holland, South-Holland, and Utrecht were chosen because these provinces are the most densely populated ones in the Netherlands, and people from these provinces consequently represent a large group of speakers.

The selected speakers have academic degrees, and jobs in accordance with their education. Although people from all social classes can be expected to speak Standard Dutch, the probability of a person speaking Standard Dutch is greater if he is well educated. As a consequence, it is easier to find speakers of Standard Dutch among the highly educated. A second reason for selecting people with academic degrees was that they would be our peers, which should positively influence the naturalness of their free conversations in our presence (§5.2.2).

5.2.4.3 Personal characteristics

The language variety spoken by a person is influenced by his sex and age. The corpus can therefore only represent similar varieties of Standard Dutch if the speakers are of the same sex, and belong to the same age group.

The choice between male and female speakers is not based on the requirements for the corpus formulated in section 5.1. Requirements (1) and (3-8) probably do not involve the speakers' sex, and it is unclear whether Requirement (2), which states that the corpus must represent Standard Dutch, is better met by male or female speakers. Sex does influence the chances of any subject speaking Standard Dutch, but studies provide contradictory observations on the nature of this influence. Brouwer (1990: 41) reports that women tend to speak more in conformity with Standard Dutch than men, whereas Stroop (1998ab) argues that it is especially women who embrace realizations which deviate from Standard Dutch.

For this reason, the choice between men's and women's speech was based on possible requirements of future research. One of these requirements may be that the corpus can be used for acoustic research. In general, men's speech is better suited to acoustic research than women's speech because men produce relatively low fundamental frequencies, which make estimates of formant values more accurate. This is why the people who were chosen as speakers for the corpus were male.

In contrast to the choice of the speakers' sex, the choice of the speakers' age could be based on the requirements of the corpus. Studies of the influence of age on language (e.g. Trudgill 1974: 104,111; Labov 1994: 47,48) show that adults' speech can be very different from adolescents' speech, and that there is also a difference between adults under or over 60. Adults between 21 and 60 years old speak more or less similarly. Since the speakers who are recorded for the corpus must speak similar varieties of Standard Dutch, it made sense to select speakers from the age group of 21-60 years old. We tried to narrow down this age group somewhat more in order to better meet Requirement (4), but this proved to be almost impossible. The restriction that one of the speakers of a pair was an acquaintance of mine, or an acquaintance's acquaintance more or less dictated an age group of 21-55, because any further restriction would have made it difficult to find enough subjects.

Finally, in order to make the corpus as representative of spoken Dutch as possible, people were chosen who had no obvious speech or hearing impediments, and spoke Dutch as their native language.

5.2.4.4 Judgements of the recorded speech

The requirement that the selected speakers should have academic degrees may increase the probability that they speak Standard Dutch, but does not guarantee it. After the recordings, a further selection process took place to ensure that only samples of Standard Dutch were included in the corpus. The classification could not be based on objective criteria, since there are no principled criteria for distinguishing Standard Dutch from dialects. That is why the recorded speech of every subject was classified on subjective grounds as Standard and non Standard by six independent judges and ourselves.

The six judges were male, had academic degrees, and sounded as speakers of Standard Dutch to us. Three judges came from the western part of the Netherlands, and were probably well qualified to distinguish western dialects from the western variety of Standard Dutch. The three other judges came from the South-East, the North-East, and the South-West of the Netherlands and were we think qualified to determine the extent to which a subject's speech is typical for the west. Appendix E.1 lists the judges' ages.

The judges listened to parts of the conversations taken from the beginning, the middle, and the end of each recording session. The exact parts were determined by the judges themselves. Their opinions of every speaker were elicited by questionnaires.

5.2.4.5 Summary

The speakers in the corpus have the characteristics listed in (1).

(1) Characteristics of the speakers.

a. They have lived in the western provinces North-Holland, South-Holland and Utrecht all their lives.
b. They have academic degrees, and jobs in accordance with their education.
c. They are male.
d. They are between 21 and 55 years old.
e. They speak Dutch as their native language.
f. They have no known speech or hearing impediments.
g. They were judged to speak Standard Dutch.

5.3 The actual recordings

Section 5.2 described the general decisions taken with respect to the design of the corpus. This section will describe the way these decisions were implemented. It will describe the actual recordings, the recorded speakers, and the classification of the recorded speech as Standard Dutch or non-Standard Dutch.

The recordings for the corpus were made between 1 November 1995 and 1 March 1996 in the soundproof room at the Institute of Phonetics of the University of Amsterdam. Several speakers with the characteristics (1a-f) (§5.2.4.5) were invited to participate in the recordings. They were told that the recordings were intended for the compilation of a corpus of Standard Dutch, the language variety that appeared to match their spontaneous speech. The ten speakers who agreed to participate were asked to bring a friend or close colleague along who, like themselves, had lived in the western part of the Netherlands all his life, and had an academic degree. The speakers were not paid for their participation, but did receive a bottle of wine afterwards.

As soon as the speakers entered the soundproof room, the recordings were invisibly started from another room. The speakers were seated at a table, at a distance of 1.5 metres from each other. Microphones of the type Sennheiser MD527 with supercardioidic direction characteristics were positioned on the table. They were placed right in front of the speakers, at a distance of 25 cm. Next to each microphone there was a potted plant, whose function was to take the speakers' minds off the microphones, and to make the room somewhat less austere, which was necessary, as the poor lighting, the foam rubber on the walls, and the grid on the floor gave the room a gloomy atmosphere. The fourth participant in the conversations, if present (§5.2.2), and the author were seated at approximately 1.5 metres from both speakers. The recordings were only stopped after the speakers had left the soundproof room.

The sounds picked up by the microphones were recorded by a DAT-recorder (Denon DTR 2000) on different tracks of a tape (Sony 90 min). Because of the distance between the speakers, a speaker sounds approximately 30 dB louder on his own track than on the other speaker's track. This makes it possible to tell the speakers apart when listening to the recordings, even when they talk simultaneously. The volume of the recording was adjusted manually and continually, with every adjustment being noted, and each recording ended with a sound of 81 dB. This means that there is a record of the intensities of all realizations, which could be useful for future research.

The speakers received oral instructions for the first free conversations, and on the careful realizations of the monosyllabic words. With respect to the part of the free conversations, the subjects were told that the three, or four, of us, would talk for approximately 35 minutes, and that the conversations had to be real conversations, and not interviews. The instructions for the second part, i.e. the role-play, were partly in written form and partly oral. The written instructions were sent to the speakers' homes at least three days before the recording session. They are the final versions of the instructions arrived at in the pilot experiments, and can be found in Appendix C. The speaker who appeared to be the most self-confident of a pair was assigned the part of the salesman, since this is the most demanding role. The other speaker was assigned the part of the store owner. Just before the speakers started playing the role-play, they received the oral instructions. They were told that they should try to be themselves as best they could, and were allowed to talk about the soundproof room, and their own occupations. They were also urged to take their roles seriously, a request which was emphasized with the promise that the winner of the negotiations would receive a bottle of wine. Finally, the speakers were instructed not to finish the role-play before being told to do so.

Table 5.2 shows the relations between the members of each pair of speakers, and whether an extra speaker was present during the first free conversations. The subjects whose speech was recorded with a view to including it in our corpus are labelled Subjects A to Q. They all posses the characteristics (1a-f). Speaker X_1's speech is not included in the corpus because this subject had a nasty cold at the day of the recording session, and consequently was often unintelligible. The speech of X_2 and X_3 is not included either because these subjects have not lived in the western part of the Netherlands all their lives. Subjects A to P all filled in the questionnaires for the study on the validity of linguistic intuitions described in section 4.2.2. Their dates of birth, places of residence, and professions can be found in Appendix A.

Table 5.2 The speakers on the recordings.

Recording	Speakers	Relation speakers	Extra speaker
I	A, B	Friends	
II	C, D	Friends	
III	E, X_1	Friends	
IV	F, G	Colleagues	
V	H, X_2	Friends	Subject X_2's sister
VI	I, X_3	Colleagues	
VII	J, Q	Friends	
VIII	K, L	Friends	
IX	M, N	Colleagues	
X	O, P	Friends	Subject O's wife

All pairs are made up of friends or direct colleagues who have approximately the same age, and are socially equal to each other. The only exception is the pair of recording I, since Subject B is Subject A's supervisor, and 19 years older. This hierarchical relationship might have created problems for the elicitation of casual speech, which is why Subject A was requested to choose another partner. However, he insisted on Subject B, and countered our objections by insisting that he spoke freely with Subject B. It is assumed here that Subject A does not feel the relationship between himself and Subject B to be hierarchical, and talks with Subject B as he does with friends.

The speech of Subjects A to Q was judged by seven other speakers of Dutch, including ourselves (§5.2.4.4). We agreed that subjects A to P can be considered speakers of Standard Dutch (see Appendix E.2 for a summary of the judgements). Since the corpus has to contain Standard Dutch (Requirement 1), the speech realized by these subjects can, and will, be used to compile it.

5.4 Evaluation

5.4.1 Introduction

Evaluating the compilation of the corpus involves answering two questions. The first one concerns the adequacy of the implementation of the design decisions (Did the recordings work out as expected?). The second question concerns the adequacy of the design itself (Did the design give rise to a corpus that meets the requirements formulated in §5.1?). Sections 5.4.2 and 5.4.3 will deal with these two questions.

5.4.2 The recordings

The recordings did not present any problems. The speakers managed to come up with enough topics to fill 90 minutes of conversation, and there are no long silences, and no long periods of laughter on the recordings. Since the speakers generally sat facing each other while they were talking, they talked well into the microphones.

The free conversations at the beginning of the recordings generally started with remarks on the gloomy atmosphere of the room, and subsequently developed very naturally. The subjects discussed widely differing topics including TV-shows, writing dissertations, sport, publishing articles, money, libraries, archaeology, children books, trams, and parking and living in Amsterdam. Two small fragments of such conversations can be found in Appendix D.

During the first part of the role-play, all pairs of subjects discussed a party, and their visits to the dentist or driving tests. In addition, they discussed some of the other conversation topics suggested in the written instructions. Additional topics which they came up with include listening to literature on tape, football matches, being locked in somewhere, types of coffee, and babysitting. Appendix D contains two small fragments of such conversations.

Several speakers spent some time gossiping. This entails that, for reasons of privacy protection, the corpus cannot be free to the public, and examples from the corpus have to be chosen with care. One speaker remarked afterwards that he and his partner gossiped so much because they thought that it was easier to speak naturally in the circumstances when gossiping.

The negotiations were generally more lively and took more time than in the pilot experiment. This was probably the case because the speakers who were recorded for the actual corpus had had more time to prepare themselves, as they had received the instructions several days before the recording session. Some pairs of speakers took so much time for the negotiations that they did not finish them on time. They were still negotiating when the 90 minutes DAT-tape was nearly full, and we had to stop the role-play. Two fragments of negotiations can be found in Appendix D.

5.4.3 The design

Section 5.1 formulated eight requirements which have to be met by the corpus. These requirements are summarized in (2) for convenience.

(2) Requirements for the corpus

1. The corpus should contain casual speech.
2. The corpus should represent Standard Dutch.
3. The subjects should speak the same variety of Standard Dutch.
4. The recordings that make up the corpus should be available.
5. The recordings that make up the corpus should be free of background noise.
6. The corpus should not be very large.
7. The corpus should be sufficiently large for the purposes of the investigations.
8. Every speaker should be represented by a relatively large number of realizations.

The corpus contains approximately 122,500 tokens of words, which corresponds to an average of 7,650 word tokens per speaker. This implies that the corpus certainly meets Requirement (6). Actual research must show whether it also meets requirements (7) and (8).

The 12 hours of recordings that make up the corpus are available, and the subjects on these recordings were classified as speakers of Standard Dutch (§5.3). The corpus therefore also certainly meets Requirements (2) and (4).

The corpus does not completely meet Requirement (5), since the speakers made background noises with their papers and plastic cups, and often spoke simultaneously. A certain percentage of tokens consequently cannot be used as data for phonological and phonetic research (cf. §8.2.4).

In order to find out whether the corpus meets Requirement (1), the six judges who evaluated each subject's language variety (§5.2.4.4) were asked for their opinion on the casualness of the recorded speech. The judges agreed with us that the subjects spoke naturally and casually, and that the corpus therefore meets Requirement (1). Some subjects realized very natural and casual speech particularly during the first free conversations, whereas others sounded more natural during the role-play, and especially during the negotiations. The original reason for the division of each recording into a part of the free conversations and a role-play was that neither part can fill 90 minutes on its own. Now it appears that the division has yet another important function: it solves a potential problem in that it ensures that both people who cannot handle role-plays very well and people who have some difficulties speaking spontaneously in more or less forced conversations in the presence of a researcher produce natural speech in some part of the recording.

For the evaluation of the corpus with respect to the compromise between Requirements (2) and (3) (§5.1), the judges and ourselves classified the subjects' varieties of Standard Dutch (see Appendix E.2). We agreed that some subjects speak plain Standard Dutch, whereas others speak a western variety of Standard Dutch, or Standard Dutch with a flavour of the dialect of Amsterdam. The judges remarked that they distinguished these varieties mainly, if not only, on the basis of the realizations of fricatives. Subjects who realized all, or nearly all, fricatives as voiceless were classified as being influenced by the dialect of Amsterdam, or as speakers of a western variety of Standard Dutch. The other subjects were classified as speakers of plain Standard Dutch. The judges did not note additional important differences between the speakers. Subject A to P therefore probably speak similar varieties of Standard Dutch, and the corpus meets both Requirements (2) and (3) as well as possible.

In conclusion, the corpus meets Requirements (1-4), and (6), whereas it does not completely meet Requirement (5). The investigations of the present study will show whether it meets Requirements (7) and (8).

5.5 The present form of the corpus

The corpus was completely transcribed orthographically. The transcription includes hesitations, repetitions, contrastive accents, and false starts. People's names are indicated with their initials in order to avoid (immediate) identification (see the parts of the transcription incorporated in Appendix D). The utterances realized by the rejected speakers (speakers X_1, X_2, X_3, and Q) are partly included in the transcriptions, because these speakers formed pairs with subjects whose speech is incorporated into the corpus, which implies that including their utterances in the transcriptions makes it easier to locate the utterances of the accepted speakers on the tapes.

The speech in the corpus is transcribed neither phonemically nor phonetically, nor does it contain much prosodic information. Providing this information is very time-consuming, and as it was not needed for all utterances, it was decided to transcribe phonetically and partly prosodically only those utterances that are relevant to the present study (see Chapters 6 and 8).

5.6 Summary

This chapter discussed the compilation of the corpus of casual Standard Dutch that forms the basis for the investigations in the following chapters. The corpus consists of approximately 122,500 tokens of words realized by 16 male subjects in roughly 12 hours of dialogues. A schematic overview of the recordings that make up the corpus is Table 5.1 in section 5.2.3. The speakers have the characteristics listed in (1) in section 5.2.4.5).

Part IV

A rough survey of phoneme realizations

6 Absent segments and reduced vowels

6.1 Introduction

The preceding two chapters dealt with the type of data that are needed for the investigations which will be carried out in this study. This chapter will present the first of these investigations, viz. the rough survey of phoneme realizations in casual Dutch. The in-depth study of the realization of obstruents as either voiced or voiceless can be found in the following chapters (Part V).

As announced in section 1.2, the survey will focus on the absence of segments which are present in highly careful speech, and on the realization of underlyingly full vowels as schwas. We will

- consider whether these non-realizations are characteristic of casual Dutch by consulting the literature on careful Dutch;
- determine the contexts in which the non-realizations are possible;
- speculate on which of the non-realizations can be completely due to the speaker's natural tendency to reduce articulatory effort;
- speculate on which other factors are relevant.

For the purposes of this survey, a data set was created consisting of transcriptions of approximately 2,500 stretches of speech representing the language of five randomly chosen subjects from the corpus described in the previous chapter, i.e. Subjects F, G, I, M and N (for more information on these subjects, see §5.2.4 and Appendix A). The transcriptions use the symbols of the International Phonetic Alphabet (International Phonetic Association 1999) including the diacritics. Vowel length is not indicated as it is considered to be irrelevant to this study, and it is often difficult to determine whether a vowel is short. Examples from the data set are given in (1).

(1) Some transcribed stretches of speech

a.

vind	*ik*	*dan*	*inderdaad*	*wel*	*mooi*
find	I	then	indeed	[modal particle]	pretty
/vɪnd	ɪk	dɑn	ˌɪndərˈdad	ʋɛl	moj/

[fɪŋktɑʔɪˈdatʋəmoj]
‘then I find pretty indeed’

b.

Maar	*ik*	*heb*	*hem*	*nooit (...)*
but	I	have	him	never (...)
/mar	ɪk	hɛb	əm	nojt/

[makɛbmnojt]
‘but I never have him (...)’

c.

probeer	*je*	*een*	*bibliotheekassistent*
try	you	a	library assistant
/proˈber	jə	ən	ˌbiblioˈtekˌɑsiˌstɛnt/

[proˈbejəmˌbijəˈtekɑsˌtɛnt]
‘one tries a library assistant’

d.

blijf	*ik*	*ongeveer*	*op*	*hetzelfde* (...)
stay	I	approximately	on	the same (...)
/blɛif	ɪk	ˌɔŋxəˈver	ɔp	hɛtˈzɛlfdə/

[blɛifkɔxˈfeɔptˈsɛldə]
‘I stay approximately at the same (...)’

In order to obtain the transcriptions, the author listened to the recordings containing the speech of the five subjects, and fed utterances which seemed relevant into the speech analysis software package *Praat* (Boersma 1996), with a sampling frequency of 48 kHz. *Praat* facilitates playing the utterances back, and is therefore a valuable listening tool. Although it can provide acoustic information, such as waveforms and intensity curves, the transcriptions were completely made by ear, since interpreting acoustic information is often difficult (§4.3.2).

On the basis of the transcriptions generalizations are formulated on the context of a certain type of realization. Every generalization finds support in at least 100 transcriptions.

Some of the stretches of speech on which these generalizations were made were not only transcribed by the author, but also by two other trained phoneticians. Each generalization was considered to be valid if all three judges arrived at identical transcriptions independently for at least one of the relevant stretches of speech. In practice, this number was often as high as five. All examples given in this chapter have been triple-checked in this way and represent unanimous judgements. The percentage of stretches on which the three phoneticians disagreed varied with the type of (non)-realization. For instance, it was low (0%) for the absence of [t], and high (over 50%) for the absence of schwa.

The contexts in which segments are absent or realized as schwa are assumed to include prosodic constituency. That is, the syllabic positions of the segments will be taken into account as well as the question whether the segments belong to stressed or unstressed syllables. Monosyllabic unaccented function words are considered to be unstressed, since they tend to incorporate into preceding feet, i.e. to be clitics (§2.3.4 and §3.6). No distinction will be made between syllables with main stress and secondary stress, and neither between syllables with different types of secondary stress, since the corpus comprises too few words with secondary stress, and can therefore provide little information on the influence of the type of stress on the (non)-realization of segments. Finally, the prosodic domains that are higher in the prosodic hierarchy than the prosodic word will be left out of consideration. The main reason for this is that spotting phrase boundaries in speech is a difficult task, and it was estimated that its costs probably outweigh the benefits for the purposes of the present study.

The data set is expected to contain mainly items of high frequency of occurrence. This expectation is based on three facts. First, high frequency items by definition occur more often in conversations. Consequently, they are more likely to be incorporated into data sets of non-systematic observations. Second, highly frequent items are more easily recognized by the listener, and therefore a speaker can hypo-articulate them to some extent without running the risk of being misunderstood (§2.3.3), which is why they surface relatively often in reduced forms. Finally, highly frequent items are expected to be overrepresented in data sets of reduced forms, because the reduced forms of highly frequent items are probably highly frequent themselves, and stored in the lexicon. Reduced realizations which are represented in the lexicon surface when they are retrieved from the lexicon, or when they result from phonological and phonetic processes. They are therefore realized more often than forms which can only result from phonological and phonetic processes.

Following the discussion in Chapter 1, we assume that reduction in articulatory effort affects the sizes of the articulatory gestures, and the overlap of the gestures in time. We hypothesize that the following types of information provide cues to what extent the speaker's natural tendency to reduce articulatory effort can result in the absence of a certain segment.

1. *The articulatory properties of the segment*
   The consequences of reduction in the sizes of the articulatory gestures depend on the regular articulatory properties of the segment at issue. Since these properties depend on the syllabic position of the segment, this position has to be taken into account (e.g. §6.2.3.2 and §6.2.9).

2. *The context of the segment*
   The consequences of reduction in articulatory effort on a certain segment depend on the articulatory and acoustic properties of the adjacent segments. For instance, whereas the complete coarticulation of [p] and [n] results in a different segment, i.e. [m], the coarticulation of [p] and [t] results in the acoustic absence of one segment, i.e. [t]. This means that the influence of the speaker's natural tendency to reduce articulatory effort on the realization of a segment can be determined only if the articulatory and acoustic characteristics of the adjacent segments are known (e.g. §6.2.2.3, and §6.2.7).

3. *The acoustic salience of the segment*
   A speaker only gives in to his natural tendency to reduce articulatory effort if the reduction does not seriously increase the probability that the message of his utterance will be misunderstood by the listener (§2.2.3). Reduction therefore typically affects segments which are not acoustically salient, i.e. short segments, segments which are relatively soft, and segments with acoustic properties very similar to those of the adjacent segments. As a consequence, the acoustic salience that a segment would have had if it had been acoustically present provides information on the probability that it has been reduced only for phonetic reasons (§6.2.2.2).

4. *The relevance of the segment for word recognition*
   Word-medial segments belonging to unstressed syllables are less relevant for recognition than initial and final segments, and segments belonging to stressed syllables (§2.2.2). Since speakers only reduce segments if reduction does not hinder communication, reduction particularly tends to affect word-medial segments of unstressed syllables. This is why these segments are more likely to fall victim to the speaker's natural tendency to reduce articulatory effort. Hence, if particularly these segments are reduced, the reduction may result from the speaker's tendency to reduce articulatory effort (e.g. §6.2.3.2, and §6.2.8).

5. *The frequency of occurrence of the item*
   Highly frequent items are more easily identified by listeners than items of low frequency (§2.2.2). Since speakers only reduce items if reduction does not obstruct communication, it is particularly items of high frequency that are expected to be affected by reduction. The frequency of an item therefore indicates the probability that some of its segments are absent on phonetic grounds (e.g. §6.2.4.4, and §6.2.6). Frequency data will be obtained from our corpus and the pilot study of the Eindhoven corpus of spoken Dutch (Uit den Boogaart 1975), which together contain approximately 244,000 tokens of words (121,569 tokens in the relevant part of the Eindhoven corpus and approximately 122,500 tokens in my corpus).

6. *Relevance of the word to the propositional content*
   Since speakers only reduce words if the reduction does not hinder communication, words which are not highly relevant to the propositional content of the utterance or are highly probable given the context are more likely to be reduced by phonetic processes than more relevant or less probable words (e.g §6.2.6). We will consider a word highly relevant if it provides new information, or provides information which is indispensable to the interpretation of the utterance. A word is highly probable given the preceding word, for instance, if it follows that word relatively often.

7. *Influence of morphological and phonological factors*
   Reduction in articulatory effort is a phonetic process. This study assumes that phonetic processes only have access to the phonological form, and therefore do not have direct access to morphological information or phonological principles (§2.4.1). This means that if the (non-) realization of a segment is influenced by morphological or phonological principles, it cannot be completely due to factors influencing phonetic implementation (e.g. §6.2.4.2 and § 6.2.5).

These types of information can at most serve to provide suggestions as to what causes the absence of a segment, not hard and fast answers. The conclusive answer to the question whether a segment is absent as a result of the speaker's natural tendency to reduce articulatory effort must come, at least partly, from articulatory data. This type of data is not available in the present study. We can therefore only speculate why segments are absent.

The chapter will first formulate generalizations on the absence of single consonants in section 6.2, and then on the absence of vowels, and the realization of vowels as schwas in section 6.3. Section 6.4 will discuss extremely reduced forms of words and combinations of words. The generalizations are illustrated with transcriptions of stretches of speech which are glossed literally.

Note that since the data set contains only a proportion of all relevant utterances, with only a minority of all transcriptions double-checked by other phoneticians, the present study cannot provide information on the frequencies of the various realizations, and generalizations that emerge from the data set, however suggestive, need to be confirmed by other, more systematic studies.

6.2 Absence of consonants

6.2.1 Introduction

When words are realized in casual speech, their acoustic forms often contain fewer consonants than when they are realized in highly careful speech. Our data set contains hundreds of stretches of speech in which [t, r, n] are perceptually absent, and dozens of stretches in which [h, x, k, d, l, f] are absent. Their absence will be discussed in sections 6.2.2 to 6.2.10. These sections will confirm observations described in previous literature, and provide new data.

6.2.2 Absence of [t]

6.2.2.1 Introduction

The absence of [t] is one of the best documented instances of reduction in the Germanic languages. It has been thoroughly studied in, for instance, English (see e.g. Guy 1980; Neu 1980), and several Dutch dialects (see e.g. de Vries 1974; Ottow-Kolman 1989; Hinkens 1992; Goeman 1999). In casual Standard Dutch it has been studied on the basis of incidental observations and intuitions (see e.g. Zwaardemaker & Eijkman 1928: 232, 233; van Haeringen 1971; Booij 1995: 152-154).

The data from the corpus show that in casual Dutch, as in careful Dutch, [t] tends to be absent particularly in the middle of consonant clusters. In this context it seems to be absent in casual Dutch especially after [s] (§6.2.2.2), and before bilabial stops (§6.2.2.3). When [t] is not part of a consonant cluster, it can be absent in casual Dutch in *niet* 'not' (§6.2.2.4), and at the end of certain verb-stems (§6.2.2.5).

6.2.2.2 Absence of [t] after [s]

When segment [t] is part of a consonant cluster, it tends not to appear in the acoustic form particularly also when it is in word-final position after [s] (see examples 2). The absence of [t] in this context does not seem to be influenced by the type of the following segment or by the presence of stress.

(2) Absence of [t] after [s]

a.	*winstmarge*	/'ʋɪnst̲-ˌmɑrʒə/	['ʋɪnsˌmɑrʒə]	'profit margin'
b.	*dienstweigeraar*	/'dinst̲-ˌʋɛixəˌrar/	['dinsˌʋɛixəˌrar]	'conscientious objector'
c.	*vast zit*	/'vɑst̲ ˌzɪt/	['fɑˌsɪt]	'sits stuck'
d.	*juist gehoord*	/'jœyst̲ xə'hord/	['jœysxə'hort]	'just heard'
e.	*daarnaast is*	*daar*/'nast̲ ɪs/	*daar*['nasɪs]	'besides it is'

The absence of [t] after [s] can be due to the speakers' wish to reduce articulatory effort, as reduction in the size of the articulatory gestures can give rise to the perceptual absence of [t] after [s]. When speakers reduce the closing gestures of a stop, the stop can be realized with the characteristics of a fricative (§1.1). When /t/ is realized as a fricative, it sounds as [s], which is almost indistinguishable from a preceding [s]. The absence of [t] after [s] may therefore result from a decrease in the size of the closing gesture.

The assumption that the absence of [t] after [s] can be due to the speaker's natural tendency to reduce articulatory effort is supported by the fact that [s] and the noise burst of [t] have the same spectral properties. Hence, a [t] following [s] is acoustically not very salient, and tends to be realized with little articulatory effort.

6.2.2.3 Absence of [t] before bilabial stops

Segment [t] is regularly absent in consonant clusters also before bilabial stops, particularly when it functions as a verbal affix (see examples 3).

(3) Absence of [t] before bilabial stops

a.	*denkt men*	/'dɛŋk-t̲ mɛn/	['dɛŋkmɛn]	'think-s one'
b.	*lijkt me*	/'lɛik-t̲ mə/	['lɛikmə]	'seem-s to me'
c.	*past best*	/'pɑs-t̲ 'bɛst/	['pɑsbɛst]	'fit-s well'
d.	*heeft besloten*	/heft̲ bə/*sloten*	[hefbə]*sloten*	'has decided'

Verbal suffixes represent a case of contextual inflection, and are therefore not highly relevant for the interpretation of Dutch sentences, which is why they could be expected to be more likely to be affected by the speaker's natural tendency to reduce articulatory effort than segments belonging to stems. The fact that it is particularly

[t] functioning as a verbal suffix that is absent before bilabial stops suggests therefore that the absence of [t] in this context is due to the speaker's natural tendency to reduce articulatory effort.

This suggestion is supported by the fact that the absence of [t] before bilabial stops may be a consequence of coarticulation. Bilabial stops are realized with a constriction at the lips, which is why they can mask all sounds that are realized within the vocal tract (Browman & Goldstein 1990a: 360). When they are realized simultaneously with [t], they can mask the release of the [t], which is the most important cue to its recognition. In such cases, the [t] is acoustically absent, although its articulatory gestures are present.

6.2.2.4 Absence of [t] in *niet*

The [t] of the word *niet* 'not', which has an underlying form /nit/, is absent in many stretches of speech in the data set. Unlike other words which end in [t], *niet* apparently tends to be realized without [t] before all types of segments, even before vowels (see examples 4).

(4) Absence of [t] in *niet* before vowels

a.	*niet aangedacht*	/nit̠ an/gedacht	[nian]*gedacht*	'not thought of'
b.	*niet echt*	/nit̠ 'ɛxt/	[niʔ'ɛxt]	'not really'
c.	*niet eens*	/nit̠ 'ens/	[ni'ens]	'not even'

The absence of [t] in *niet* cannot solely be due to the speaker's natural tendency to reduce articulation effort, since neither reduction in the size of the articulatory gestures nor coarticulation can account for the complete absence of [t] of *niet* in all segmental contexts. The highly frequent realization [ni] suggests that *niet* has the lexical variant /ni/.

6.2.2.5 Absence of [t] in certain verb-forms

In addition to *niet*, certain present tense singular verb forms, such as *weet* 'know', *vind* 'find', *moet* 'must', and *word* 'is' ('passive auxiliary'), appear without their final alveolar stop in many stretches in the data set. Their stops can apparently be absent in the acoustic forms of the second person singular when these forms are followed by the corresponding subject pronoun *je* /jə/ or *jij* /jɛi/ 'you', and in the acoustic forms of the first person singular before all types of segments (see examples 5). The stops are underlyingly voiced or voiceless, but, when realized, they are generally voiceless, as they are in coda position (§3.4.2).

(5) Absence of stem-final alveolar stops

a.	*vind ik*	/'vɪnd̲ ɪk/	['fɪnɪk]	'am of the opinion I'
b.	*ik vind het*	*ik* /'vɪnd̲ ət/	*ik* ['fɪnət]	'I think it'
c.	*vind je*	/'vɪnd̲ jə/	['fɪjə]	'are of the opinion you'
d.	*moet je*	/'mut̲ jə/	['mujə]	'must you'
e.	*ik weet niet*	*ik* /'ʋet̲ nit/	*ik* ['ʋeni]	'I know not'

The regular present-tense verb forms in Dutch have the forms indicated in (6).

(6) Regular present-tense verb forms

first person singular:	*verb stem*
second person singular	
- when followed by the subject pronoun:	*verb stem*
- else:	*verb stem* + t
third person singular:	*verb stem* + t
plural:	*verb stem* + ən

When the suffix /t/ follows a /d/ or /t/, the resulting sequence of two alveolar stops is realized as [t]. The present-tense forms of two verbs are given in (7).

(7) Present tense forms of two verbs

a. verb stem: /lop/ 'walk'

first person singular:		*loop*	/lop/	[lop]
second person singular:		*loop*	/lop/	[lop]
	or	*loopt*	/lop-t/	[lopt]
third person singular:		*loopt*	/lop-t/	[lopt]
plural:		*lopen*	/'lop-ən/	['lopən]

b. verb stem: /vɪnd/ 'be of the opinion'

first person singular:		*vind*	/vɪnd/	[vɪnt]
second person singular:		*vind*	/vɪnd/	[vɪnt]
	or	*vindt*	/vɪnd-t/	[vɪnt]
third person singular:		*vindt*	/vɪnd-t/	[vɪnt]
plural:		*vinden*	/'vɪnd-ən/	['vɪndən]

Since the alveolar stop is absent in the forms of the first person singular, and in the forms of the second person singular which are followed by the corresponding subject pronoun, it is typically absent when the verb stem is not followed by a suffix. That is, it is absent when it corresponds to the final segment of the verb stem.

Stem-final /t/ is part of the lexical representation of a word, unlike suffixal /t/, which is a case of contextual inflection. Therefore, stem-final /t/ is more important for the recognition of a word, and hence for communication, than suffixal /t/. If reduction in effort is responsible for the absence of [t] in the forms of the relevant verbs, suffixal [t] would be expected to be absent more often than stem-final [t] (§6.2.2.3). This is contrary to fact, and therefore the absence of the alveolar stop in the cases at issue is probably not solely due to the speaker's natural tendency to reduce articulatory effort.

A more likely explanation is the one proposed by Van Haeringen (1971), who assumes that stem-final [t] is often absent because it is erroneously interpreted as a verbal suffix, and therefore deletes, as the relevant suffix should be ∅.

Van Haeringen (1971) also assumes that highly frequent stems are realized more often without their final alveolar stops than stems of low frequency. This assumption cannot be tested on the basis of the data set, as this set consists of non-systematic observations, and therefore does not contain all tokens of all verb forms ending in an alveolar stop in the corpus.

If van Haeringen's assumption is correct, a possible explanation is that the individual forms of highly frequent verbs are generally retrieved as wholes from the lexicon, whereas the forms of other verbs are usually computed from their parts (§2.2.1). Segments in the verb forms which are retrieved as units are probably not explicitly marked as either part of the verb stem or a suffix, since these forms are more or less regarded as single units. In contrast, the segments which represent suffixes in the other verb forms, i.e. the ones that are computed from their parts, are probably marked as suffixes. Hence, it is easier to erroneously interpret a stem-final alveolar stop as a suffix when it is part of a verb form that is retrieved as a single unit from the lexicon than when it is part of a computed verb form. The alveolar stops in the former verb forms then are more likely to be absent, and the resulting forms without alveolar stops are probably stored as variants in the lexicon. These forms, which are the shorter ones, are retrieved especially in casual speech.

6.2.2.6 Summary

To recapitulate, the data suggest that all types of words which are realized with [t] in careful speech have acoustic forms without [t] in casual speech. The [t] is absent mainly when it follows [s], or precedes bilabial stops. The absence of [t] in these contexts could be completely due to the speaker's natural tendency to reduce articulatory effort. In addition, the [t] is absent particularly in the word *niet*, and in the highly frequent verbal forms which consist only of a verb stem ending in an alveolar stop. The absence of [t] in these forms probably has to be ascribed to the fact that the lexicon contains reduced variants of words and highly frequent complex words.

6.2.3 Absence of [r]

6.2.3.1 Introduction

The phoneme /r/ has many variants in Dutch. Which one is realized depends e.g. on the syllabic position of the segment, and the geographical background of the speaker (see e.g. Gussenhoven & Broeders 1976: 127; Collins & Mees 1981: 168-171; Vieregge & Broeders 1993; van Reenen 1994; Voortman 1994: 116-121; Van de Velde 1996; Van de Velde & van Hout 1999). The data set based on the corpus contains hundreds of transcriptions in which an underlyingly present /r/ is missing. These transcriptions show that coda [r] is absent especially after schwa, and between low vowels and consonants, while onset [r] is absent primarily in the word *precies* 'precise'/ 'precisely'. Missing [r] has not been reported yet for careful Dutch.

6.2.3.2 Absence of coda [r] after schwa

The absence of coda [r] after schwa (see examples 8) could be due to the speaker's natural tendency to reduce articulatory effort for two reasons. First, the [r] is in the coda of an unstressed syllable, and therefore tends to be realized with little articulatory effort. Second, if it is realized after schwa at all, it is acoustically not salient, as it tends to be realized as [ə] itself: its presence is signalled at best by the length of the vowel resulting from the sum of the two schwas. The result is that the speaker may be tempted not to realize [r] after schwa at all.

(8) Absence of coda [r] after schwa

a.	*anders*	/'ɑndər̠s/	['ɑndəs]	'different'
b.	*ouders*	/'ɑudər̠s/	['ɑudəs]	'parents'
c.	*over de*	/'ovər̠ də/	['ofədə]	'over the'
d.	*beter kan*	/'betər̠ kɑn/	['betəkɑn]	'better can'
e.	*zonder alle*	/'zɔndər̠ 'ɑlə/	['zɔndə'ɑlə]	'without all'

6.2.3.3 Absence of coda [r] after low vowels

The absence of coda [r] between a low vowel and a consonant (see examples 9) could similarly be due to the speaker's natural tendency to reduce articulatory effort. When [r] is realized after low vowels, it generally sounds like a short [j] type of glide (Collins & Mees 1981: 170). For the realization of this sound, the mouth has to be slightly closed. This closing gesture may coincide with the gesture needed for the articulation of the following consonant, in which case the [r] is not perceptible, even if phonologically present. Consequently, the absence of [r] after low vowels and before consonants may be due completely to coarticulation.

It is particularly the function words *daar* and *waar* that were found without

[r].The reason for this might be that they are highly frequent (§6.1). The word *daar* has approximately 1678 tokens in our corpus and the pilot study of the Eindhoven corpus of spoken Dutch, and the morpheme *waar* 485.

(9) Absence of coda [r] after low vowels

a.	*daarna*	/dar̠ˈna/	[daˈna]	'after that'
b.	*waar je*	/ʋar̠ jə/	[ʋajə]	'where you'
c.	*waarschijnlijk*	/ʋar̠ˈsxɛinlək/	[ʋaˈsxɛinlək]	'probably'

The data do not show that the absence of [r] after low vowels is influenced by stress.

6.2.3.4 Absence of onset [r] in *precies*

When /r/s is not realized in onset position, it is generally part of the word *precies* /prəˈsis/ 'precise' / 'precisely'. The first syllable of this word is variously realized as [prə], [pər], [pə], or [p].

The absence of [r] in *precies* is possibly due to coarticulation of [p] and [r]. The [p] is realized at the lips, and can therefore hide [r], which is generally realized within the vocal tract. The stop [p] is particularly likely to hide the sonorant [r], since stops tend to be much longer than sonorants (Waals 1999: 22).

Another possibility may be that the absence of [r] in *precies* results from the high ranking of a phonological constraint which requires onset consonant clusters to be as simple as possible. This possibility is supported by the realization [pərsis], which obeys this constraint, but nevertheless contains the [r].

Onset [r] was probably found to be absent in the word *precies* because this word seldom bears accent and is not a low frequency word, as it has 192 occurrences in the two corpora. It can be hypo-articulated somewhat, and may have a reduced lexical variant.

6.2.3.5 Summary

In conclusion, coda [r] is principally absent after schwas and low vowels, while onset [r] is primarily absent when it is part of the word *precies.* The absence of coda [r] may be due to the speaker's natural tendency to reduce articulatory effort. The absence of [r] in *precies* could also be due to the high ranking of a phonological wellformedness constraint.

6.2.4 Absence of [n]

6.2.4.1 Introduction

The [n] is another consonant of Dutch that can be absent in the acoustic forms of words realized in casual speech. Its absence is more general and better perceptible than, for instance, the absence of [r], and was noticed by many linguists (e.g. Blanquaert 1924, 1934; Booij 1995). The data set contains hundreds of utterances in which [n] is imperceptible. They indicate that [n] tends to be absent particularly after schwa, before obstruents within words, and in word-final position.

6.2.4.2 Absence of [n] after schwa

In careful as well as casual Dutch [n] may be absent after schwa. It may be absent for instance, in *regen* /rexən/ [rexə] 'rain', *Nijmegen* [nɛimexə] 'Nijmegen', and *eten* /et-ən/ [etə] 'eat-inf.'.

According to Koefoed (1979) and Booij (1995: 139), [n] can be phonetically absent after schwa only if it does not precede unstressed vowel-initial suffixes, and is not part of the morpheme *een* /ən/ 'a', or a verbal stem. This implies that the absence of [n] is directly conditioned by morphological factors, and cannot be completely due to factors which influence phonetic implementation. The absence of [n] after schwa must result from phonological constraint interaction.

Many studies, including several quantitative ones have been devoted to the absence of [n] after schwa in careful and more or less casual Dutch (e.g. Koefoed 1979; van Oss & Gussenhoven 1984; Van de Velde 1996). A rough survey of the non-realization of [n] in this context that is based on non-systematic observations from the corpus is therefore unlikely to expand our knowledge, which is why we will restrict ourselves to the observation that in the corpus [n] is often absent after schwa too.

6.2.4.3 Absence of word-medial [n] after full vowels

Blanquaert (1924: 75, 76), Gussenhoven & Broeders (1976: 140) and others claim that word-medial [n] after all types of vowels can be absent before consonants in careful Dutch, and that its absence generally cooccurs with a nasalized realization of the preceding vowel. The non-systematic observations from the data set show that this is possible in casual Dutch as well (see examples 10), but that the following consonant is nearly always a fricative (cf. van Reenen & Coetzee 1996: 76).

(10) Absence of word-medial [n] before consonants. The preceding vowels are nasalized (not indicated in the transcriptions).

a.	*mensen*	/'mɛn̠sən/	['mɛsə]	'people'
b.	*tenminste*	/tɛn'mɪn̠stə/	[tə'mɪstə]	'at least'
c.	*ongeveer*	/ˌɔn̠xə'ver/	[ˌɔxə'ver]	'approximately'
d.	*informeren*	/ˌɪn̠fɔr'merən/	[ˌɪfɔ'merə]	'to inform'

The absence of word-medial [n] before fricatives can be explained by the speaker's natural tendency to reduce articulatory effort. Word-medial [n]s are realized with the same place of articulation as the immediately following fricative (Booij 1995: 64). Hence, a word-medial [n]-fricative sequence is realized with a constriction at only a single position in the vocal tract. This constriction has to be complete during the realization of the [n], and released somewhat for the realization of the fricative. If a speaker does not take the trouble to form the complete constriction first, the [n] sounds like the following fricative, and the two phones are almost indistinguishable, with the [n] appearing to be absent. The perceptual absence of word-medial [n] before fricatives can therefore result from a reduction in articulatory effort.

If it is true that even when [n] is acoustically absent its gestures are still present (in a reduced form), the nasalized realization of the preceding vowel could similarly be due to the speaker's tendency to reduce articulatory effort. Instead of opening his velum at a high speed for the realization of the nasal consonant just at the beginning of this consonant, the speaker could choose to do this slowly during the realization of the preceding vowel. The vowel is then nasal as a result of coarticulation with the nasal consonant.

The assumption that the vowel is nasal as a result of coarticulation is not adopted by van Reenen & Coetzee (1996: 78). They state that a vowel preceding a nasal consonant is generally realized with a greater velic opening than the nasal consonant, and is nasal in its own right: its nasality is not the result of articulatory coarticulation, but of perceptual assimilation. Because no gestural data on casual Dutch are available, it cannot be determined whether van Reenen & Coetzee are right.

6.2.4.4 Absence of word-final [n] after full vowels

The absence of [n] in word-final position after full vowels seems to be restricted to a limited number of word types (see examples 11). These words are realized without [n] mainly before word-initial fricatives and sonorants, and sometimes before word-initial stops (11a-e). Before word-initial vowels, [n] can apparently be absent only when it belongs to the word *dan* /dɑn/ 'then' (11f), which suggests that this word has a lexical variant without /n/. The absence of [n] generally cooccurs with a nasal realization of the preceding vowel.

(11) Absence of the word-final [n] of function words. The preceding vowels are nasalized (not indicated in the transcriptions).

a.	*en van*	/ɛn̲/ *van*	[ɛ]*van*	'and of'
b.	*gaan ze*	/xan̲/ *ze*	[xa]*ze*	'go they'
c.	*kan veel*	/kɑn̲/ *veel*	[kɑ]*veel*	'can much'
d.	*gaan regelen*	/xan̲/ *regelen*	[xa]*regelen*	'go and arrange'
e.	*toen ben*	/tun̲/ *ben*	[tu]*ben*	'then am'
f.	*en toen*	/ɛn̲/ *toen*	[ɛ]*toen*	'and then'
g.	*dan ook*	/dɑn̲ ok/	[dɑʔok]	'then also'

The absence of word-final [n] after full vowels may well have the same source as the absence of word-medial [n] after full vowels, i.e. reduction in the size of the articulatory gestures (see above). If this is the case, it has to be explained why [n] is more likely to be absent before non-fricatives when it is word-final than when it is word-medial. A possible explanation is the influence of frequency of occurrence, since the word-medial [n]s in the corpus are all part of low frequency words, whereas the word-final [n]s are part of high frequency words. As mentioned above, the influence of frequency can be ascribed to the fact that listeners recognize words of high frequency more easily, and that more high frequency words have reduced lexical variants.

6.2.4.5 Summary

In conclusion, /n/ may have no correspondent in the acoustic form when it closes a syllable headed by schwa, or follows a (nasalized) full vowel while preceding a consonant, especially a fricative. The absence of [n] after schwa is generally phonological in nature. Its absence after full vowels possibly results from the speaker's natural tendency to reduce articulatory effort.

6.2.5 Absence of [h]

The absence of [h] has not yet been reported for careful Dutch. In casual Dutch it is not as general as the absence of [t, r, n]. The data show that [h] is especially absent in forms of the verb *hebben* /hɛb-ən/ 'have-inf.' which follow a consonant and do not bear accent (see example 12a).

When [h] is absent in a form of *hebben* which follows one of the schwa-final subject pronouns, viz. *we* /ʋə/ 'we', *zij* /zə/ 'she' or 'they', or *je* /jə/ 'you', the preceding schwa is generally absent as well (see examples 12bc), so that the absence of [h] does not result in an onsetless syllable.

(12) Absence of [h]

a.	*ik heb*	/ɪk h̲ɛb/	[ɪkɛp]	'I have'
b.	*ze hadden*	/zə 'h̲ɑd-ən/	['sɑdən]	'they had'
c.	*we hebben*	/ʋə 'h̲ɛb-ən/	['ʋɛbə]	'we have'

Segment [h] was also absent in the data set in an acoustic form of the compound *achterhoofd* /ɑxtər-hofd/ 'back (of the) head', which has the careful realization [ɑxtərhoft]. Subject M realized this word as ['ɑxtroft], i.e. without [h] and without [ə]. This phonetic form appears to contain the syllables (ɑx) and (trof).[1] If the [h] had been present, the segments [tr], which form a syllable with the schwa in careful realizations, could not have been part of the same syllable as [of], since [h] does not allow other segments in the onset: [(trhof)$_\sigma$] is not a phonologically wellformed syllable. Hence, the absence of [h] allows for a wellformed prosodic structure without [ə].

The syllabification [(ɑx)$_\sigma$(trof)$_\sigma$t] suggests that the morphological constituents *achter* and *hoofd* do not form prosodic words of their own. If they did, a syllable boundary would have separated the [r] and the [o], as these two segments would then belong to different prosodic words. Apparently, the part *hoofd* is incorporated into the prosodic word of *achter*, while maintaining its foot structure.

In conclusion, the absence of [h] appears to be restricted to a limited number of words, and often coincides with adjustments to prosodic constituency. Since the prosodic structure of a string is determined in phonology, the interaction between the presence of [h] and prosodic constituency indicates that the absence of [h] results from phonological constraint interaction.

6.2.6 Absence of [x]

The data set shows that in casual speech [x] can be absent in the acoustic forms of the function words *nog* /nɔx/ 'still' and *toch* /tɔx/ 'yet', especially when these forms precede labial consonants or [n] (see examples 13). The absence of [x] in this context has not as yet been reported for careful Dutch.

The fricative [x] is realized by a constriction near the velum, while labials and [n] are realized by a constriction at the lips, or near the alveolar ridge. Hence, coarticulation of a [x] with a labial or a [n] may make the [x] nearly inaudible. This

---

[1] The [t] forms an appendix and is directly adjoined to the prosodic word (§3.6).

implies that the acoustic absence of [x] before labials and [n] may well result from the retiming of gestures, i.e. the speaker's natural tendency to reduce articulatory effort.

(13) Absence of [x]

a.	*nog wat*	/nɔx̠ ʋɑt/	[nɔʋɑt]	'still something'
b.	*nog niks*	/nɔx̠ nɪks/	[nɔnɪks]	'still nothing'
c.	*toch best*	/tɔx̠ 'bɛst/	[tɔ'bɛst]	'yet okay'
d.	*toch niet*	/tɔx̠ nit/	[tɔnit]	'yet not'

The assumption that the absence of [x] results from reduction in articulatory effort is confirmed by the observation that it is mainly restricted to *nog* and *toch*. These words are highly frequent, as they are represented by 1400 and 1008 tokens in the two corpora, and are not highly relevant for communication. Speakers can hypo-articulate them to a certain extent without running the risk of being misunderstood.

If the absence of [x] before [n] and labials is indeed due to coarticulation, [x] may be expected to be frequently absent before other alveolars as well, such as [t] and [d]. The data set does not support this hypothesis. A possible explanation is that tokens of *nog* and *toch* almost invariably precede negative words, which are always [n]-initial. This means that data sets of non-systematic observations are far more likely to contain tokens of these words without [x] before [n] than before other consonants. Moreover, it means that word-combinations of *nog* and *toch* plus a following negation are possibly stored in the lexicon, and that they have a lexical representation with /x/ and one without /x/. As a consequence, [x] is absent in particular before negation words, and therefore before [n].

6.2.7 Absence of [k]

Whereas [k] seems never to be absent in careful Dutch, it can be absent in careful speech. Several stretches of speech in the data set show that in casual Dutch [k] can be absent if it is the final segment of *denk* /dɛŋk/ 'think'. For the realization of an [ŋk] cluster, a constriction has to be formed at the soft palate, and the velum must first be open for the realization of the [ŋ] and then be closed for the [k]. If the velum is not yet closed when the constriction at the soft palate is released, the /ŋk/ sequence sounds as [ŋ]. Hence, the absence of [k] after [ŋ] can be the result of the retiming of gestures, and be completely due to the speaker's natural tendency to reduce articulatory effort. Tokens of *denk* are particularly likely to be realized without [k], possibly because this word is highly frequent.

6.2.8 Absence of [d]

In contrary to careful Dutch, in casual Dutch [d] can be absent between [n] and [ə] (see examples 14). Like an [ŋk]-cluster, an [nd]-cluster can be reduced to its nasal by retiming of the relevant gestures: the oral stop is perceptually absent when the alveolar constriction which is needed for both consonants is released before, instead of after, the velum is closed. The absence of the [d] of an [nd]-sequence can therefore be due to the speaker's natural tendency to reduce articulatory effort. Hypo-articulation is encouraged since the /d/ is part of an unstressed syllable.

(14) Absence of [d]

a.	*anders*	/'ɑnd̲ərs/	['ɑnərs]	'different'
b.	*inderdaad*	/ˌɪnd̲ər'dad/	[ˌɪnər'dat]	'indeed'
c.	*indertijd*	/ˌɪnd̲ər/*tijd*	[ˌɪnər]*tijd*	'at the time'

6.2.9 Absence of [l]

In casual Dutch, [l] tends to be absent particularly in unaccented tokens of the function word *als* /ɑls/ 'like' / 'if'. If realized, the [l] is in coda position in this word, and is therefore generally realized as vocalic (see Collins & Mees 1981: 166; van Reenen 1986), i.e. as a colouring of the preceding vowel. The perceptibility of the colouring depends on the length of the vowel. If the vowel is acoustically short, the articulators have almost no time to change positions, and the colouring is hardly perceptible. The vowel is short in most tokens of *als* for two reasons. First, *als* is highly frequent — it has 1528 occurrences in the two corpora — and highly frequent words tend to be shorter than words of low frequency with the same number of phonemes. Second, *als* is generally unstressed, and unstressed vowels tend to be shorter than stressed ones (Nooteboom 1972: 75). Hence, the /l/ of *als* is often almost imperceptible and appears to be absent. This implies that the perceptual absence of [l] in *als* does not necessarily result from the absence of [l] in the phonological form. It further implies that speakers may choose not to realize the [l] of *als* at all, since listeners are aware of the fact that it is often acoustically absent. In other words, the realization [ɑs] could be due to the speaker's natural tendency to reduce articulatory effort.

Since the realization [ɑs] is highly frequent, it is probably stored in the lexicon. The absence of [l] can therefore probably also be lexical in nature.

6.2.10 Absence of [f]

The last consonant which is absent in dozens of tokens in the data set is [f]. It is absent particularly in tokens of the words *zelfde* /ˈzɛlfdə/ 'same' and *zelfs* /zɛlfs/ 'even'. Its absence has not as yet been reported for careful Dutch

The absence of [f] in *zelfde* and *zelfs* can result from a decrease in the size of the labial gesture. The segments surrounding [f] in both words are alveolar, and are therefore probably coarticulated, i.e. realized with a single alveolar constriction. The bilabial gesture needed for the realization of [f] should be realized at the top of this alveolar constriction, and only serves the realization of the [f]. This may explain the reduction in the size of this gesture. The words *zelfde* and *zelfs* allow hypo-articulation, as they seldom bear stress.

6.2.11 Conclusions

Words are sometimes realized with fewer consonants in casual speech than in highly careful speech. The absence of most consonants can be due to coarticulation, or reduction in the sizes of the articulatory gestures, i.e. to the speaker's natural tendency to reduce articulatory effort. The absence of consonants in some words is probably encoded in the lexicon.

Whether the absence of a particular consonant results from the speaker's natural tendency to reduce articulatory effort, from phonological constraint interaction, or is lexical in nature cannot be conclusively determined on the basis of the audible properties of speech, but requires psycholinguistic data as well as data on the gestures involved in the reduced realizations, neither of which are available.

6.3 The realization of vowels

6.3.1 Introduction

It is not only the realizations of consonants, but also those of vowels that are affected in casual speech. In casual speech,

- all types of vowels are often absent (e.g. Dalby 1984; Kohler 1990);
- full vowels are often realized as schwas (e.g. Kager 1989);
- voiced vowels are sometimes realized as voiceless (Ohala 1983: 203; Rodgers 1998);
- tense vowels sometimes sound as their lax counterparts (Nooteboom 1979: 258).

When the phoneticians transcribed utterances from the corpus, it appeared that they disagreed in many cases on the [voice]-characteristics of vowels, and on the difference between [a] and [ɑ], and [o] and [ɔ]. This is why the present study does not deal with the realization of voiced vowels as voiceless, and tense vowels as lax, but focuses instead on the realization of full vowels as schwas (§6.3.2), and the absence of vowels (§6.3.3). The realization of full vowels as schwas provides information on the prosodic structure of compounds and word-combinations in casual speech (§6.3.4).

6.3.2 The realization of full vowels as schwas

The frequency with which an underlyingly full vowel in a word is realized as schwa in Dutch is claimed to be influenced by the factors in (15).

(15) Factors influencing the realization of vowels as schwas

a. *Presence of stress*
Stressed vowels tend not to be realized as schwas (Rietveld & Koopmans-van Beinum 1987; van Bergem 1993, 1995).

b. *Presence of an onset*
Vowels in onsetless syllables are seldom realized as schwas (Martin 1968: 167; Booij 1981: 148).

c. *Quality of the vowel*
The lower the position of a vowel in the following hierarchy, the more often it is realized as schwa: /y, u, ø/ > /i/ > /o, ɔ/ > /a, ɑ/ > /e, ɪ/. Diphthongs are hardly ever realized as schwas (Stroop 1974: 320; Kager 1989: 300; Booij 1995: 131, 134).

d. *Presence of a coda*
Vowels in open syllables are more often realized as schwas than vowels in closed syllables (Booij 1995: 131).

e. *Position in the word*
Vowels tend not to be realized as schwas if they belong to word-final syllables (Kager 1989: 304; Booij 1995: 131).

f. *Frequency of occurrence*
Vowels are more often realized as schwas if they are part of high frequency words than of low frequency ones (Booij 1995: 130; van Bergem 1995).

As all studies of vowel reduction to schwa are only based on linguistic intuitions, which we claimed to be invalid for non-lexical processes (§4.2), or on sentences read aloud, it would be interesting to see whether there is any support for the factors listed in (15) in our data set of realizations in spontaneous conversations.

The factors in (15) were investigated on the basis of hundreds of non-systematic observations from our corpus. Vowels are considered to be schwas if they sound as being located in the centre of the vowel space, or as having inherited the place specifications of the surrounding segments (§3.2). Vowels that are underlyingly specified as /ʏ/ are left out of consideration, as their full forms are almost impossible to distinguish from the reduced ones.

The data appear to support the generalizations that the presence of stress and onsets influence the realizations of vowels as schwa, since no reduced vowel in the data set is contained in a stressed or onsetless syllable. Generalizations (15ab) are apparently in line with data from casual speech.

With respect to the influences of factors (15c-f), the data are less conclusive. As to (15c), the data show that diphthongs can be realized as schwas in casual Dutch (see examples 16). The diphthong /ɛi/ turned out to be reduced especially in the highly frequent word *altijd* /ɑltɛ̱i̱d/ 'always' realized with stress on the first syllable (see 16bcd), and in *bij* /bɛ̱i̱/ 'with' followed by a pronoun (16e). This implies that (15c) is not obviously correct for casual Dutch, and that only systematic research can confirm or falsify the claim that diphthongs are realized as schwas less often than monophthongs in this speech style.

(16) /ɛi/ realized as schwa

a.	*aardrijkskunde*	/'ardrɛiks,kʏndə/	['arəs,kʏndə]	'geography'
b.	*altijd prima*	/'ɑltɛid 'prima/	['ɑltət'prima]	'always excellent'
c.	*altijd op*	/'ɑltɛid ɔp/	['ɑltət ɔp]	'always on'
d.	*bij mij*	/bɛi 'mɛi/	[bə 'mɛi]	'with me'

No conclusions can be drawn with respect to the influence of the presence of coda consonants (15d) and the position of the vowel in the word (15e), since the data comprise several closed and word-final syllables with reduced vowels (see examples 17), which shows that reduced vowels are possible in these types of syllables in casual speech. Generalizations (15de) are therefore not obviously correct for casual speech, and should be tested in systematic quantitative research.

(17) Closed and word-final syllables with reduced vowels

a.	*bijvoorbeeld*	*bijvoor*/belt/	*bijvoor*[bəlt]	'for instance'
b.	*allemaal*	/'ɑlə,mal/	['ɑməl]	'all'
c.	*altijd*	/'ɑltɛit/	['ɑltət]	'always'

This conclusion also holds although vowels are reduced particularly in syllables that are realized without their final consonants (see examples 18), that is, vowels which are in closed syllables at the lexical level of phonology are more likely to be reduced when their syllables are realized as open. This latter observation does not necessarily support (15d) as it could simply be due to the fact that both the realization of vowels as schwas and the absence of some consonants are favoured in unstressed syllables.

(18) Reduced vowels in syllables realized without their coda consonants

a.	*waarschijnlijk*	/ʋar'sxɛinlək/	[ʋə'sxɛinlək]	'probably'
b.	*verleden*	/vɛr'ledən/	[fə'ledə]	'past'
c.	*dat weet ik*	/dɑt/ *weet ik*	[də] *weet ik*	'that know I'
d.	*dan word*	/dɑn/ *word*	[də] *word*	'then become'

Finally, the data do not provide conclusive evidence for the hypothesis that the realization of vowels as schwas is favoured in highly frequent words (15f), in spite of the fact that most words in the data set are highly frequent. High frequency words are more likely to be incorporated into the data set, and the fact that they have more tokens with reduced vowels in absolute terms in the data set is therefore meaningless. Only systematic quantitative studies which take into account the relation between the number of occurrences of the words in the corpus and in the data set can determine whether there is a difference between high frequency words and rare words.

In conclusion, the non-systematic observations from the corpus suggest that all types of vowels which are unstressed and follow onsets can be realized as schwas in casual speech. They present no evidence that the frequency with which a vowel is realized as schwa is influenced by the underlying quality of the vowel, the presence of a coda, the position of the vowel in the word, or the frequency of occurrence of the word. The influences of these factors can only be evaluated in systematic quantitative research.

The realization of full vowels as schwa implies a reduction in articulatory effort, since schwas are generally realized with a shape of the vocal tract that requires the least articulatory effort under the circumstances (cf. §3.2). Nevertheless, vowel reduction is apparently a phonological process, since it is sensitive to the presence of onsets.

6.3.3 Absence of vowels

6.3.3.1 Introduction

The absence of unstressed vowels is an important characteristic of the casual speech styles of Germanic languages (see e.g. Dalby 1984; Kohler 1990). The data from the corpus indicate that unstressed vowels which are present in formal Dutch are absent in casual Dutch particularly in three types of contexts: in vowel hiatus position, between obstruents and liquids, and adjacent to continuants. The absence of vowels often coincides with the absence of adjacent consonants.

6.3.3.2 Absence of vowels in hiatus position

Whereas in careful Dutch unstressed word-initial and word-final [ɪ] and [ə] are generally perceptible, they are regularly imperceptible in casual Dutch if they are separated from another vowel by a grammatical word boundary only (see examples 19). If [ə] and [ɪ] happen to be adjacent, it is schwa which is optionally absent (see examples 19de). The absence of [ə] and [ɪ] in this context avoids vowel-hiatus within the prosodic word, and is apparently phonological in nature.

(19) Absence of [ɪ] or [ə] adjacent to another vowel within the same prosodic word

a.	*bij het*	/bɛi ət/	[ˈbɛit]	'at the'
b.	*sta ik*	/ˈsta ɪk/	[ˈstak]	'stand I'
c.	*hij is*	/ˈhɛi ɪs/	[ˈhɛis]	'he is'
d.	*samen in*	/ˈsamən ɪn/	[ˈsamɪn]	'together in'
e.	*dagen in*	/ˈdax-ən ɪn/	[ˈdaxɪn]	'day-s in'
f.	*rugzakken en*	*rug*/ˌzɑk-ən ɛn/	*rug*[ˌsɑkɛn]	'backpack-s and'

Examples such as (19) show that function words as well as content words can be realized without their final [ə]. In the majority of cases, the vowels of function words are absent, because there are no content words starting with schwa, and only few starting with [ɪ]. Moreover, there are no words ending in [ɪ], and most content words ending in schwas are plural or infinitive forms, which are seldom followed by vowel-initial words within the same prosodic word.

6.3.3.3 Absence of schwa between obstruents and liquids

A schwa can be absent in casual speech also when it follows an obstruent and precedes a liquid and a unstressed vowel within the same prosodic word (see examples 20). It can be absent in this context also in careful speech (see Booij 1995: 128-130), and the relevant process is therefore probably phonological in nature.

(20) Absence of schwa between an obstruent and a liquid followed by a unstressed vowel within the same prosodic word

a.	*makkelijk*	/ˈmɑkələk/	[ˈmɑklək]	'easy'
b.	*vreselijk*	/ˈvresələk/	[ˈfreslək]	'terrible'
c.	*later het*	/ˈlatər ət/	[ˈlatrət]	'later it'
d.	*beter een*	/ˈbetər ən/	[ˈbetrən]	'better a'

6.3.3.4 Absence of vowels adjacent to continuants

All types of unstressed vowels can be absent in casual Dutch when they would follow or precede continuants, especially voiceless fricatives. This is exemplified in (21). Note that the schwa of *kijken* in (21f) would be next to the continuant [ʋ], because the [n] is not realized.

(21) Absence of unstressed vowels adjacent to continuants

a.	*vorige*	/'vɔrɪxə/	['vɔrxə]	'previous'
b.	*verschillende*	/vɛr'sxɪl/*ende*	[f'sxɪl]*ende*	'several'
c.	*bijvoorbeeld*	/bɛi'vor/beeld	['bvor]beeld	'example'
d.	*vervelende*	*ver*/'veləndə/	*ver*['feIndə]	'annoying'
e.	*helemaal*	/'heləmal/	['helmal]	'completely'
f.	*kijken we*	/'kɛik-ən ʋə/	['kɛikʋə]	'look-pl. we'

The absence of vowels in this type of context is impossible in careful Dutch. Probably, it is phonetic in nature, as it sometimes leads to phonologically ill-formed syllables in the acoustic forms. The sequences [fsx] and [bv] in (21bc), for instance, are not possible onsets in Dutch, and the [n] in (21d) can neither form a well-formed coda with the preceding [l], nor a well-formed syllable on its own, nor a well-formed onset with the following [d] (Booij 1995: 24, 39, 40). Under the assumption that segments must be part of well-formed syllables in the phonology of a language, the absence of the vowels must consequently be due to phonetic processes. Further support for this can be found in the fact that it is often impossible to classify the vowels as absent or very short. The vowels may apparently be fully absent, fully present, or something in between, which implies that their presence is gradual rather than categorical, and that their absence is due to phonetic processes.

If the absence of full vowels adjacent to continuants is phonetic in nature, the absence of these vowels is possibly not exclusively the result of extreme reduction in their duration, but also of coarticulation with the adjacent continuant. Coarticulation can explain why vowels are absent particularly when adjacent to voiceless fricatives: since such fricatives are generally realized without glottal vibration, vowels partly coarticulated with these fricatives also tend to be realized without glottal vibration, in which case they are voiceless, and cannot be perceived.

6.3.3.5 Absence of vowels plus adjacent consonants

The absence of vowels in the contexts described in sections 6.3.3.3 and 6.3.3.4 tends to coincide with the absence of adjacent consonants. This is at least partly the case because vowels and some consonants can be absent under the same conditions, i.e. when part of unstressed syllables. In addition, consonants and vowels are sometimes simultaneously absent because several types of consonants are almost impossible to be realized or perceived when not preceded or followed by vowels. Coda /r/s are a case in point. Coda /r/ is often realized as a mere colouring of the preceding vowel. If this vowel is absent, the /r/ is consequently absent as well (see 22ab). Two other examples are the prefixes *be-* /bə/, which is completely absent especially before /d/ or a bilabial stop, and *ge-* /xə/, which is absent especially before /x/ (see e.g. 22cde).

The absence of the vowels of these prefixes before [d, p, b] and [x] results in [bd], [bp], [b:], and [x:]. These sequences are almost indistinguishable from [d], [p], [b] and [x], which are the second members of these clusters. Therefore, when the vowels of these prefixes are absent before [d, p, b] and [x], the preceding obstruents generally appear to be absent as well.

(22) Absence of unstressed vowels and adjacent consonants

a.	*uiterlijk*	/'œytərl/*ijk*	['œytl]*ijk*	'at the very least'
b.	*over te* verb	/'ovər tə / verb	['oftə] verb	'again to *verb*'
c.	*bepaalde*	/bə'paldə/	['paldə]	'certain'
d.	*bedoel*	/bə'dul/	['dul]	'mean (verb)'
e.	*gegevens*	/xə'xevəns/	['xevəs]	'data'

6.3.3.6 Summary

To recapitulate, there are three types of context in which vowels are optionally absent in casual Dutch. First, unstressed /ɪ/ and /ə/ at the beginnings and endings of grammatical words can be absent when they are in hiatus position within prosodic words. Second, schwas can be absent when they are preceded by an obstruent and followed by a liquid and a unstressed vowel. Third, unstressed vowels adjacent to continuants are optionally absent. The absence of vowels in the first and third context appears to be restricted to casual speech.

The absence of vowels in the first two types of context is probably phonological in nature, whereas the absence of vowels next to continuants is probably due to purely phonetic causes. Since schwas before liquids come under both the second and third context, they can be absent both for phonological and phonetic reasons.

6.3.4 The prosodic structure of strings of words

We saw in section 6.3.2 that reduced vowels occur in unstressed syllables with onsets. If it is assumed that reduced vowels must occur in such syllables, certain realizations in the data set confirm the assumption that prosodic words may correspond to several constituents of a morphological compound, or several grammatical words (§3.6). We will discuss three examples. The discussion adopts the assumptions on prosodic constituency in Dutch discussed in section 3.6. These assumptions are, however, not crucial for the point that is made.

The first example shows that a prosodic word may correspond to several constituents of a morphological compound. It concerns the compound *aardrijkskunde* /ardrɛikskʏndə/ 'geography', which consists of the constituents *aardrijk* 'earth', the linking sound *s*, and *kunde* 'knowledge'. The part *aardrijk* is a compound itself, and consists of the lexical morphemes *aard*(*e*) 'earth', and *rijk* 'domain'. The whole compound *aardrijkskunde* is probably stored as a unit in the lexicon, since it typically refers to a subject of secondary school, a fact which cannot be predicted on the basis of its parts. Because *aardrijkskunde* contains three lexical morphemes, i.e. three morphological parts, it may be expected to consist of three prosodic words. The realization [arəskʏndə], which has some tokens in the data set, suggests that this is not always the case. Under the assumption that reduced vowels occur only in unstressed syllables, the realization of *rijks* with a schwa indicates that *rijks* must be unstressed, and therefore does not form a foot or prosodic word on its own in [arəskʏndə], but with the preceding or following part of the compound. Probably, it forms a foot and prosodic word with *aard*, since *aardijks* was realized as [arəs], which forms a trochee, the perfect foot in Dutch.[2] Figure 6.1 shows this prosodic structure of [arəskʏndə]. Apparently, the form [arəskʏndə] contains fewer prosodic words than lexical morphemes.

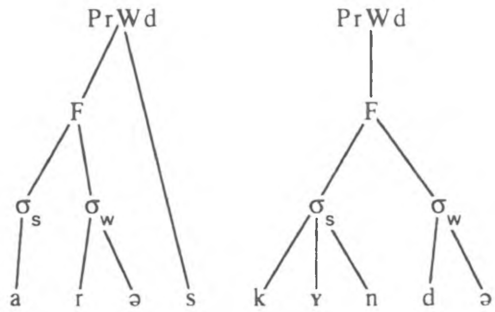


Figure 6.1 The prosodic structure of [arəskʏndə].

The second example shows that a single prosodic word may correspond to several grammatical words. It concerns the expressions *oud en nieuw* /ɑud ɛn niw/ and *af en toe* /ɑf ɛn tu/. These expressions must be stored in the lexicon, as their meanings cannot be derived from their parts: *oud en nieuw* 'old and new' is a name for the turn

2 This conclusion is further supported by the fact that *aardrijk* also forms a single prosodic word in the frequent realization [ardrɛikskʏndə] for *aardrijkskunde*. Here, the /d/ of *aard* is realized as [d], instead of [t], and therefore must be in onset position, i.e. form a syllable, and a prosodic word, with *rijk*.

of the year, and *af en toe* 'off and to' means actually 'now and then'. The corpus contains several tokens of the two expressions in which the word *en* was realized with a schwa. Under the assumption that reduced vowels must be part of syllables with onsets and without stress, these tokens show that the vowel of *en* is not always part of a stressed syllable without onset in these expressions. It sometimes forms an unstressed syllable with the preceding consonant, i.e. [d] or [f]. Being unstressed, this syllable cannot be the head of a prosodic word. Apparently, the realizations of *oud en nieuw* and *af en toe* with schwa, i.e. [ɑutəniw] and [ɑfəntu], have the prosodic structures given in Figures 6.2 and 6.3. These structures support Booij's claim that unstressed function words can be incorporated in the prosodic word headed by the preceding content word (§3.6).

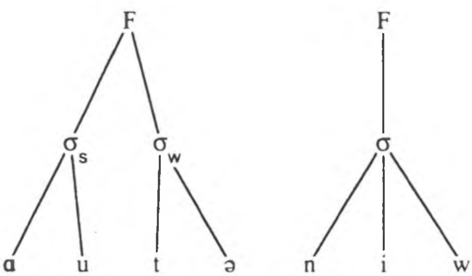


Figure 6.2 The prosodic structure of [ɑutəniw].

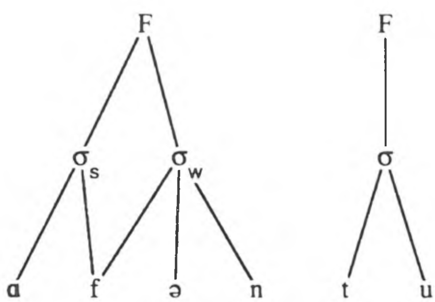


Figure 6.3 The prosodic structure of [ɑfəntu].

The final example also shows that a prosodic word may correspond to several grammatical words. It concerns the highly frequent, and therefore probably lexically stored, phrases *niet alleen* /nit ɑ'len/ 'not only' and *een keer per* /en ker pɛr/ 'one time a'. These expressions are generally realized with accent on *alleen* and *een*. If all grammatical words form prosodic words on their own, none of the vowels of these phrases can be realized as schwas, since they are all stressed or belong to syllables without onsets. The realizations [ˌnitə'len] and ['ekəpə], which are represented by several tokens in the data set, show that the /ɑ/ of *alleen*, the /e/ of *keer*, and the /ɛ/ of *per* can be realized as schwas. This means that they can be unstressed, i.e. contained in the dependent syllables of feet, and be preceded by onsets. When the /ɑ/ of *alleen* is realized as schwa — and therefore unstressed and preceded by an onset — it forms an unstressed syllable with the preceding [t] of *niet* and the following [l] $((\text{tɑl})_\sigma)$. This syllable forms a foot with the preceding stressed syllable $(\text{ni})_\sigma$, since Dutch feet consist of a stressed syllable followed by unstressed ones. Hence, *niet* and *alleen* form one single prosodic word (see Figure 6.4).

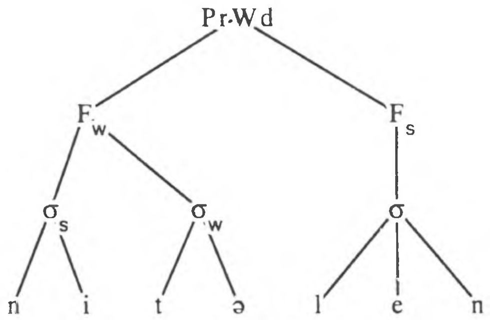


Figure 6.4 The prosodic structure of [ˌnitə'len].

When *keer per* is realized as [kəpə], the string probably forms a foot, and a single prosodic word, with the preceding stressed syllable *een* (see Figure 6.5).

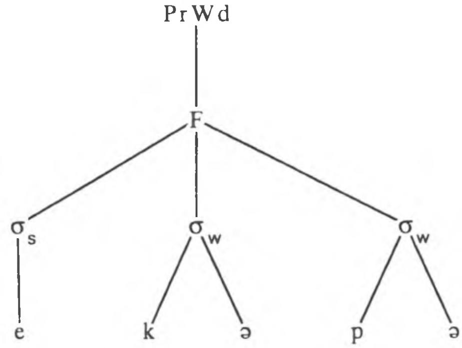


Figure 6.5 The prosodic structure of ['ekəpə].

Both prosodic structures are unexpected. The word *alleen* is incorporated into the preceding prosodic word while retaining its stress pattern, an incorporation which has been observed so far only for morphological compounds (cf. §2.2.3), and *keer* is part of a prosodic word that is not headed by the stressed syllable of the only content word, but by the syllable of the function word which normally bears accent.

In conclusion, the examples confirm the assumptions that the constituents of compounds do not always form prosodic words of their own, and that an unaccented grammatical word can be incorporated into the prosodic word of the preceding word. In addition, they suggest that the head of a prosodic word corresponding to several grammatical words is the syllable which normally bears accent.

The relevant examples are probably all retrieved as units from the lexicon. They therefore suggest that storage in the lexicon influences the prosodic structure, and consequently the realization, of morphological compounds and word-combinations.

6.3.5 Conclusions

Vowels can be realized as schwas and be completely absent in casual Dutch. As in careful Dutch, the realization of full vowels as schwas is disfavoured in stressed and onsetless syllables. The absence of vowels seems to be restricted to three types of contexts: in hiatus position within prosodic words, between obstruents and liquids, and next to continuants. The absence of vowels in the first and third type of context is characteristic for casual speech. The absence in the first and second type probably results from phonological processes, whereas the absence in the third type is probably due to the speaker's natural tendency to reduce articulatory effort. The realization of full vowels as schwas in Dutch confirms the assumption that prosodic words can correspond to several lexical morphemes in Dutch, and suggests that storage influences realization.

6.4 Extremely reduced forms

6.4.1. General aspects

The previous sections mainly dealt with the absence of single segments. In casual Dutch, unlike careful Dutch, most words and word-combinations, however, appear with several of their segments missing. Some relevant examples are listed in (23).

(23) Items realized without several of their segments

a	*op een gegeven moment*	/ɔp ən xə'xevən mo'mɛnt/	[ˌɔpə'xemənt]	'at a certain point'
b.	*volgend*	/'vɔlxənd/	['fɔlnt]	'next'
c.	*volgens mij*	/'vɔlxəns 'mɛi/	['fɔls'mɛi]	'according to me'
d.	*allemaal*	/'ɑləˌmal/	['ɑməl]	'all'
e.	*helemaal*	/'heləˌmal/	['hemal]	'completely'

The maximally reduced forms which were attested for some other items are listed in (24).

(24) Maximally reduced forms

a.	*daarom*	/'darɔm/	['dam]	'therefore'
b.	*mogelijk*	/'moxələk/	['mok]	'possible'
c.	*waarschijnlijk*	/ʋar'sxɛinlək/	[ʋ'sxɛik]	'probably'
d.	*allemaal*	/'ɑləˌmal/	['ɑmɑl]	'all'
e.	*ongeveer*	/ˌɔnxə'ver/	[ɔ'fer]	'approximately'
f.	*in ieder geval*	/ɪn 'idər xə'vɑl/	['ifɑl]	'in any case'

Examples (24a-e) indicate that the acoustic form of an item can consist merely of its initial and final segments, and the segments in the onsets and nuclei of the stressed syllables. The form (24f) is at variance with this generalization, since it does not contain the initial segment of the underlying form.

The observation that the initial and final segments of the underlying forms and the onsets and nuclei of the stressed syllables are present in nearly all maximally reduced forms fits in with the assumption that initial, final, and stressed segments are highly important for the recognition of items (§2.2.2). They probably have to be present in extremely reduced forms since otherwise recognition of the relevant items is extremely hard, if not impossible. These segments can only be absent when there are many cues to aid recognition.

This assumption implies that the codas of stressed syllables should also be present in maximally reduced forms, since they are important for recognition as well (§2.2.2). This is contrary to fact (see 24ce). Probably, the codas of stressed syllables are not always present because they cannot be identified as well as onsets: they are not followed by vowels and therefore cannot be identified by the following formant transitions, as can onsets. As a consequence, they are probably not as relevant to the recognition of items as the initial and final segments, and the onsets and nuclei of the stressed syllables, and can be left out.

If the only constraint that has to be satisfied by all reduced forms is that the items can be recognized, there is no reason to assume that the absence of non-highly relevant segments is determined in phonology, as their absence may well be due to the speaker's natural tendency to reduce articulatory effort.

Exception (24f) can then be explained as follows. The form does not contain the first segment, which is an /ɪ/, and the stressed syllable, which is /i/, because this would result in the sequence [ɪi]. This sequence is almost impossible to pronounce or perceive in fast speech, and is therefore reduced to one of its parts. The reason why it is realized as [i] may be that [i] is a stressed vowel, and ensures that the reduced form is not interpreted as *in geval* /ɪn xəvɑl/ (Vincent van Heuven, personal communication).

An alternative account of the data is that maximally reduced forms ~~result from~~ the ranking of certain phonological wellformedness constraints above faithfulness constraints. Pater (1997) claims that this assumption explains child truncations. He argues that child truncations result from the high ranking of three constraints: (i) a constraint requiring that a word should consist of maximally one foot, (ii) a constraint stating that stressed and final elements are always realized, and (iii) a constraint requiring that onsets should have the smallest possible degree of sonority. The high ranking of the third constraint has the effect that the great majority of word-initial obstruents replace the more sonorous onsets of the stressed syllables, and are retained in the truncations.

Under the assumption that the shapes of the maximally reduced forms in casual speech are determined by approximately the same phonological constraints as the child truncations, form (24f) is no exception. The first segment of the underlying form is then absent in the acoustic form because it is an unstressed vowel. The form (24c) [ʊsxɛik] for /ʊarsxɛinlək/ is then an exception, as it contains the initial segment of the underlying form, although this segment does not form a well-formed onset.

This latter account of the maximally reduced forms faces a problem. It implies that maximally reduced forms are phonological in nature, and therefore contrast with the intermediate forms which must be analysed as partly resulting from phonetic processes (§6.2 and §6.3). This is unexpected, since the intermediate reduced forms discussed in this chapter, and the intermediate forms in (23) in particular, suggest that the acoustic forms of an item form a kind of continuum between the full form and the maximally reduced form. One would not expect the forms on a continuum to have completely different sources.

Section 6.4.2 will provide additional evidence for the assumption that the acoustic forms of an item form a continuum, and that these forms partly result from phonetic processes. This section focusses on the words *natuurlijk* 'of course' and *eigenlijk* 'in fact', which are highly frequent, and were realized with various numbers of segments in the corpus.

6.4.2 *Natuurlijk* and *eigenlijk*

The words *natuurlijk* /natyrlək/ 'of course' and *eigenlijk* /ɛixənlək/ 'in fact' have 748 and 431 tokens in our corpus and the pilot study of the Eindhoven corpus of spoken Dutch.

The word *natuurlijk* consists of the lexical morpheme *natuur* /natyr/ 'nature' and the adjective/adverb-forming suffix *-lijk* /lək/. Nevertheless, most of its tokens mean 'of course' (see examples 25ab), or function as discourse markers with a meaning such as 'as you know' (see example 25c).

(25) Examples of *natuurlijk* meaning 'of course' or functioning as a discourse marker, with free translations.

a. Subject L — *Ja want daar wordt gespeeld,*
Yes, they play music there,

en dat is <u>natuurlijk</u> altijd ontzettend leuk.
and <u>of course</u> that's always immensely amusing.

b. Subject G — *B. in Leiden rekent ook uit hoeveel bibliotheken (...)*
B. in Leiden also calculates how many libraries (...)

een boek aanschaffen en dan maken ze hun calculaties
will purchase a book and then they make their calculations

zodanig dat ze quitte spelen met
in such a way that they break even with

bibliotheekprijzen over de wereld
the prices of the libraries all over the world

en (...) daar draaien ze nog steeds op,
and (...) they are still running on that,

hoewel B. was natuurlijk toch de laatste jaren
although the last years B was [discourse marker]

wat moeilijker financieel.
somewhat more difficult financially.

The word *eigenlijk* means 'in fact' (see e.g. Geerts & Heestermans 1992), but is often used as a discourse marker meaning something like 'now I am thinking about it' (see examples 26).

(26) Examples of *eigenlijk* functioning as a discourse marker. The sentences are translated freely.

a. Subject D *Ik heb er een vreselijke hekel aan, aan al dat Engels*
I detest that, all that English

dat er ingebakken zit en ik haat het eigenlijk
that is ingrained in it and I hate the fact [discourse maker]

dat ik daar de Nederlandse alternatieven niet voor ken.
that I don't know the Dutch alternatives for it.

b. Subject F *(...) dan krijg je van die losse draadjes en die*
(...) this results in those loose little threads and they

komen tussen je rits en daar heb ik eigenlijk
get stuck in your zip and I received [discourse marker]

een paar keer klachten gehad.
some complaints about it several times.

In order to investigate the types of realizations of *natuurlijk* and *eigenlijk*, all the tokens of these words in the corpus were transcribed by the author and two trained phoneticians. We incorporated a token into the data set if we arrived at the same transcription independently. The only differences we accepted in the transcriptions of a token were the difference between [a] and [ʌ] for the first vowel of *natuurlijk*, differences in the exact quality of the first vowel of *eigenlijk*, and differences in the exact place of articulation and the [voice]-characteristics of the fricative in *eigenlijk*. The tokens appeared to represent 14 types of realization for *natuurlijk*, and 6 for *eigenlijk*, which are listed in (27).

(27) The types of realization of *natuurlijk* and *eigenlijk*

a. *natuurlijk* /natyrlək/

[na'tyrlək], [nə'tyrlək], [na'tylək], [n^{ə}'tylək], [n'tylək], [n'tyl^{ə}k], [nə'ty^{l}k], [nə'tyk], [n^{ə}'tyk], [n'tyk], [n'dyk],
['tylək], ['tyl^{ə}k], ['tyk].

b. *eigenlijk* /ɛixənlək/

['ʔɛixələk], ['ʔɛixlək], ['ʔɛixlk], ['ʔɛixək], ['ʔɛixk], ['ʔɛik]

The forms of *natuurlijk* which show no trace of [na] probably are phonetic forms of /tyrlək/, instead of /natyrlək/. There are two types of evidence suggesting this lexical form. First, the realization [tyrlək] sometimes surfaces in careful speech, and in utterances consisting only of *natuurlijk*. This is unexpected if /tyrlək/ is not a lexical form, since the non-realization of /na/ is an extreme reduction, and extremely reduced forms seldom appear in careful speech, or in accented positions. Second, there are only a small number of tokens that were transcribed with an initial [n] by one transcriber and without [n] by the two others, or vice versa: we disagreed on the presence of [n] in only 20 out of 361 cases. This small percentage of disagreement (i.e. 5.5%) suggests that /na/ is either realized or not, and there is no continuum between its presence and absence in the phonetic form. Its absence therefore does not result from phonetic processes, but is a matter of phonology or the lexicon. Since there are no other words which can surface without their initial syllable [na], assuming a phonological process that results in the non-realization of [na] would make the grammar too powerful. We therefore assume that the absence of [na] is not phonological, but a matter of the lexicon, and that *natuurlijk* has the lexical variant /tyrlək/. Booij (1999: 67) argues that the variant [tyrlək] is more well-formed in Dutch than [natyrlək], since it consists of a trochee, the perfect Dutch foot. This wellformedness of [tyrlək] possibly explains why it is highly frequent, and stored in the lexicon.

Under the assumption that the [t]-initial tokens of *natuurlijk* have the underlying form /tyrlək/, all forms in (27) contain at least the initial and final segments, and the onsets and nuclei of the stressed syllables of the underlying forms. They seem to form a continuum between the full form and the maximally reduced form, since it is possible to distinguish at least the subhierarchies in (28). Each form in (28) differs from the adjacent forms in the realization of only one segment, and the forms higher in a hierarchy are more reduced than those lower in that hierarchy.

(28) Subhierarchies of realizations of *natuurlijk* and *eigenlijk*

a. *natuurlijk*

[na'tylək] > [nᵊ'tylək] > [n'tylək];

[nə'tyˡk] > [nə'tyk] > [n'tyk];

['tylək] > ['tylᵊk].

b. *eigenlijk*

['ʔɛixələk] > ['ʔɛixlək] > ['ʔɛixlk] > ['ʔɛixk] > ['ʔɛik].

The overall continuum must have more than one dimension, since the realizations do not allow a neat ranking from less reduced to more reduced. For instance, the ranking of [nətyk] with respect to [ntylək] is impossible. The different dimensions result from the fact that the segments do not all depend on each other for their realizations. The realization of /na/ of *natuurlijk*, for instance, does not seem to be strongly related to the realization of /lək/.

Some intermediate realizations are partly due to phonetic processes. The forms which were realized with a short vowel or no vowel at all corresponding to the /a/ of /na/ of *natuurlijk* are cases in point. There were 274 tokens of *natuurlijk* that were realized with [n], and therefore 274 tokens that could have been realized with a vowel after the [n]. Our transcriptions agreed only in 116 (i.e. in 42% of) cases on whether a vowel was present or absent. This high percentage of disagreements indicates that the difference between the presence of the vowel and its absence was almost impossible to perceive. Its absence seems, therefore, to result from a gradual phonetic process of vowel shortening or coarticulation, and not from a phonological constraint interaction, which has by definition categorical results. This is also indicated by the many transcriptions in which the vowel is transcribed as very short.

In conclusion, the realizations of *natuurlijk* and *eigenlijk* constitute another piece of evidence for the assumption that there is a continuum between the full form of an item and its maximally reduced form. The forms which are intermediate partly result from phonetic processes.

Since all forms on a continuum would be expected to result from the same type of process, the maximally reduced forms are not expected to be due completely to phonological constraint interaction. This implies that the first account offered in section 6.4.1 of the maximally reduced forms is more plausible than the second one. The maximally reduced forms probably result from phonological and phonetic processes which affect all segments which are not highly relevant for recognition. The continuum of reduced forms reflects the compromises between ease of articulation and ease of perception.

In addition to phonological and phonetic segmental processes, prosody and storage probably play a role in the realizations of *natuurlijk* and *eigenlijk*. The influence of prosody appears from the fact that the subjects generally realize the suffix /lək/ as [k] only in the middle of Intonational Phrases. The influence of the lexicon is suggested by Subject O's realizations. This subject is the only one who realized *natuurlijk* as [tyk] in any position of an Intonational Phrase. He realized [tyk] even when there were no other words in the utterance. His exceptional behaviour can be explained by assuming that he has the form /tyk/ stored in his lexicon, and is therefore freer in the use of this variant. This subject may possess the lexical variant /tyk/, in contrast to all other subjects, because he utters *natuurlijk* more often: he realized it 109 times, whereas each of the other subjects realized the word maximally 24 times while their speech was being taped.

6.5 Conclusions

This chapter discussed the absence of consonants and vowels and the realization of full vowels as schwas in casual Dutch. The discussion was based on a large number of transcribed stretches of speech from the corpus. Since its purpose was to give a short overview of the contexts in which these (non-)realizations are possible, it was not based on systematic research.

It appeared that segments are absent especially in contexts in which their absence can be due to reduction in the size of the relevant articulatory gestures and gestural overlap, and does not prevent the listener from interpreting the sentence. Segments tend to be absent particularly when they are part of highly frequent words, which suggests that these words have reduced lexical variants.

The study of realized vowels confirms the assumptions that vowels can be realized as schwas when they are part of unaccented syllables with onsets, and that prosodic words can correspond to several lexical morphemes. The lexical morphemes which together form single prosodic words suggest that prosodic structure is influenced by storage.

Finally, the study of maximally reduced forms of items suggests that these forms consist of the segments which are highly relevant to recognition. Probably, the maximally reduced forms and the intermediate reduced forms form a continuum which reflects the compromise between ease of articulation and ease of perception.

Future studies should involve articulatory data, as only these data can provide conclusive evidence for the assumption that the absence of several types of segments results from the speaker's natural tendency to reduce articulatory effort. These studies should also be more systematic, so that they can provide data on the frequencies of occurrence of the reduced realizations.

Part V

The realization of obstruents as voiced or voiceless

7 A new analysis

7.1 Introduction

Section 3.3 showed that the difference between voiced and voiceless obstruents is phonemic in Dutch. This difference is regularly neutralized, since the realization of any type of obstruent in word-final position or syllable-final position and the realization of some types of obstruent in morpheme-initial position is not determined by the underlying [voice]-specification of the obstruent, but by the type of adjacent segments (§3.4).

This chapter will propose a new analysis of [voice] in Dutch. This analysis assigns an important role to the speaker's tendency to reduce articulatory effort, and, in contrast to previous analyses, can explain all available data. The validity of the analysis will be tested on the basis of our corpus in the following chapters. The investigation will provide new data on Dutch, suggesting that the speaker's natural tendency to reduce articulatory effort and the exact content of the lexicon influence realizations in casual speech. These data also suggest that the analysis is an improvement over earlier analyses.

This chapter is structured as follows. Section 7.2 will discuss the problems of previous analyses of [voice] in Dutch.[1] These problems show that we need a new analysis which assigns an important role to phonetics. Section 7.3 will extensively discuss the acoustic characteristics of voiced and voiceless obstruents. This discussion forms a necessary background for the remainder of the chapter, and for chapters 8 to 10. Finally, sections 7.4 and 7.5 will propose the new analysis.

In this and the following chapters, the following terms have the following meanings:

---

[1] We will focus on the problems which beset both the classical generative analyses as the analyses framed within Optimality Theory. For an extensive discussion of the problems of the classical generative analyses, see Ernestus (forthcoming).

• *coda obstruents*:	obstruents which are only in coda position as well as obstruents which are in appendices.
• *faithful obstruents*:	obstruents which are realized in accordance with their underlying [voice]-specifications.
• *neutral obstruents*:	coda obstruents and word-final obstruents. These obstruents are called neutral since the realization of these obstruents is not related to their underlying [voice]-specifications. The distinction between underlying voiced and voiceless obstruents is neutralized for these obstruents.
• *onset obstruents*:	obstruents which are completely in onset position, or are ambisyllabic.
• *unspecified obstruents*:	obstruents which are unspecified for [voice].
• *release:*	release of an articulatory constriction.
• *voiced obstruents*:	obstruents which are perceived as voiced.
• *voiceless obstruents*:	obstruents which are perceived as voiceless.

7.2 Problems of previous analyses

Section 3.4 discussed several analyses of [voice] in Dutch. Nearly all of them assume that neutral obstruents are always realized as voiced before voiced stops, and as voiceless in all other contexts (§3.4.2). In addition, they nearly all assume that the realization of all obstruents is determined in phonology, i.e. that obstruents which are realized as voiced are phonologically specified as [+voice], and that those which are realized as voiceless are phonologically specified as [−voice]. Finally, some analyses account for certain realizations by assuming specific prosodic structures.

The assumption that word-final obstruents (e.g. Trommelen & Zonneveld 1979; Booij 1981, 1995; Berendsen 1983, 1986; Zonneveld 1983; Lombardi 1995b, 1999) are always realized as voiced before voiced stops, and as voiceless before all other types of consonants is not in line with the available data. Several descriptive studies have shown that word-final obstruents are sometimes realized as voiced before sonorants (Zwaardemaker & Eijkman 1928: 226; van Rijnbach & Kramer 1939; Gussenhoven & Bremmer 1983), and as voiceless before voiced stops (e.g. Kaiser 1958; Demeulemeester 1962; Slis 1982, 1983).

The assumption that the realization of all obstruents is determined in phonology is problematic as well, as its adoption does not provide an adequate account for the influences of the factors in (1) on the realization of word-final obstruents before voiced stops.

(1) Factors influencing the realizations of word-final stops before voiced stops

a. *Type of preceding segment*
Word-final obstruents before voiced stops are realized as voiced less often when they follow voiceless segments than when they follow voiced ones (Demeulemeester 1962). For instance, the /t/ in the word-combination *leest bovendien* /lest bovəndin/ 'reads moreover' would be less often realized as voiced than the /t/ in the word-combination *eet bovendien* /et bovəndin/ 'eat(s) moreover'.

b. *Presence of stress*
Word-final obstruents are more likely to be realized as voiced when the following underlyingly voiced stops are part of stressed syllables (Slis 1983). For instance, the obstruent cluster in *staatsbeheer* /stats+bəher/ 'state management' without stress on the second syllable would be less often realized as completely voiced than the cluster in *staatsblad* /stats+blɑd/ 'Statute-Book' with stress on the second syllable.

c. *Speech rate*
When speech rate is high,
- clusters consisting of a word-final stop and an underlyingly voiced stop are more often realized as completely voiced;
- clusters consisting of a word-final fricative and an underlyingly voiced stop are more often realized as either completely voiced or voiceless;
than when speech rate is low (Kaiser 1958; Slis 1982; Menert 1994).

d. *Speaker's sex*
Men are more likely to realize word-final obstruents before voiced stops as voiced than women (Kaiser 1958; Slis 1982, 1983).

e. *Speaker's mood*
When speakers are emotional, they are less likely to realize word-final obstruents before voiced stops as voiced (van Ginneken 1935; Meinsma 1958; Demeulemeester 1962).

Under the assumption that phonology completely determines the realizations of neutral obstruents, the influences of factors (1cde) imply that phonological rules, or the rankings of phonological constraints, differ with speech rate, speaker gender, and speaker mood. The influences of the type of preceding segment and stress on the realization of the obstruent (1ab), in addition to the influence of the following type of segment, imply complex rules or constraint interactions. These implications are undesirable.

Finally, the assumption that the realization of obstruents is influenced by prosodic structure above the foot-level can be problematic, as Gussenhoven's (1986) analysis illustrates. Gussenhoven explains the voiced realization of post-vocalic obstruents before some vowel-initial words by means of assumptions about the presence of prosodic word (PrWd) boundaries: he assumes that these obstruents can be realized as voiced when they are directly followed by PrWd boundaries (§3.4.3). In addition, Gussenhoven explains the observation that the initial /d/ of many function words, unlike the initial /d/ of content words, can be realized as voiceless after obstruents by means of assumptions about the presence of PrWd boundaries. He assumes that content words are always preceded by PrWd boundaries, and that their initial /d/s are therefore always realized as voiced. In contrast, function words are not always preceded by PrWd boundaries. If they are not, but are directly preceded by obstruents, their initial /d/s are realized as voiceless (§3.4.5). There are two reasons why Gussenhoven's analysis is problematic. First, the majority of the PrWd boundaries he assumes are not supported by independent evidence. Second, his analysis cannot well account for the observation that intervocalic obstruents before *ik* /ɪk/ 'I' are often realized as voiced, while the initial /d/s of function words are often realized as voiceless. According to Gussenhoven's analysis, this observation implies that *ik* is rarely completely incorporated into the preceding PrWd, unlike /d/-initial function words. This is an undesirable result, since there is no reason to assume that the pronoun *ik* is less clitic-like than /d/-initial function words such as *daar* and *dat*.

In summary, phonological analyses of the realization of obstruents as voiced or voiceless are not adequate. An alternative analysis is needed which does not assume that the realization of final obstruents is completely determined in phonology, and makes no unsupported claims with respect to the presence of PrWd boundaries.

We will propose such an analysis in sections 7.4 and 7.5. First, we will discuss the acoustic properties of voiced and voiceless obstruents in section 7.3.

7.3 The phonetics of voiced and voiceless obstruents

7.3.1 Introduction

We saw above that an adequate analysis of the realization of obstruents as voiced or voiceless in Dutch cannot be purely phonological in nature. Since phonology and phonetics are the components which determine the realization of segments, this implies that such an analysis must take phonetics into account. The formulation of an adequate analysis, then, should be based on our knowledge of the phonatory, articulatory, and acoustic properties of voiced and voiceless obstruents. Sections 7.3.2 and 7.3.3 will discuss general phonetic properties of stops and fricatives, while section 7.3.4 will deal with the difference between strings containing voiced and voiceless obstruents.

7.3.2 Stops

A speaker produces a stop by forming a complete constriction somewhere in the vocal tract, and then releasing this constriction rapidly. During the constriction phase, no sound is emitted from the vocal tract, or at most a weak sound produced by the vibration of the vocal folds (e.g. Ladefoged & Maddieson 1996). The release of the constriction normally produces a noise burst, the spectral properties of which are important cues for the place of articulation of the stop (e.g. Cooper et al. 1952; Halle et al. 1957). The movements of the articulators to and from the constriction of an intervocalic stop result in rapid changes in the formant frequencies. These changes are called transitions, and are also cues to the place of articulation of the stop (Cooper et al. 1952; Halle et al. 1957).

7.3.3 Fricatives

A fricative is a friction noise, i.e. an air turbulence caused by friction. Speakers realize fricatives by forming a very precisely shaped channel between two articulators, and pushing air into this channel (see e.g. Ladefoged & Maddieson 1996). The sound results from the turbulence generated by friction at the constriction ("tunnel turbulence"), or, in the case of sibilants, from the turbulence generated when the jet of air formed at the constriction strikes the edge of some obstruction, such as the teeth ("wake turbulence").

7.3.4 Acoustic cues to the voiced/voiceless distinction

Languages vary in the acoustic cues for the voiced/voiceless distinction. In Dutch, the difference can be signalled by the following cues.

1. *Presence of vocal fold vibration*

   Voiced obstruents tend to be realized with vocal fold vibration during the last part of the constriction, whereas voiceless obstruents do not. During the realization of a voiced stop in intervocalic position the vocal folds often vibrate continuously (Lisker & Abramson 1964, among others).

   For the realization of vocal fold vibration (see e.g. Clark & Yallop 1995: 187; Rietveld & van Heuven 1997: 60), the glottis has to be closed, and the lungs have to press air on the vocal folds. As long as the air pressure below the glottis is approximately 2-4 cm H_2O higher than the air pressure above the glottis, there is a net pressure from below and the glottis is opened, and a small amount of air is pushed into the vocal tract. Since this air is pushed through a narrow opening, it accelerates. The acceleration of the airflow leads to a drop in the air pressure between the vocal folds (the Bernoulli effect), and a closure of the glottis. The glottis is opened again by the net pressure from below if the difference in air pressure above and below the glottis is approximately 2-4 cm H_2O.

   The realization of a fricative with vocal fold vibration is difficult. The air pressure above the glottis has to be lower than the pressure beneath it for the realization of vocal fold vibration, and at the same time, it has to be higher than the pressure beyond the constriction for the realization of friction. These two requirements can be met simultaneously with undue effort only for a short time (Ohala 1983).

   The realization of a long oral stop with vocal fold vibration is difficult too. During the constriction of an oral stop, no air can escape from the vocal tract. This implies that during glottal vibration the air pressure above the glottis increases every time some air passes the glottis. At a certain moment, the pressure above the glottis is so high that the difference between the pressure above and below the glottis drops below 2-4 cm H_2O, and glottal vibration stops. Speakers may delay this moment somewhat by expanding their vocal tract, i.e. by decreasing the supraglottal air pressure. They move the glottis downward, and move the tongue root forward (Westbury 1983).

   The sounds produced by glottal vibration during acoustically short intervocalic obstruents are drowned out by the sounds of the preceding vowels.

As a consequence, the presence/absence of vocal fold vibration during these obstruents is difficult to perceive, and cannot be an important cue for the voiced/voiceless distinction (Slis & van Nierop 1970).

2. *Acoustic duration of the obstruent*
Voiceless fricatives are generally longer than voiced fricatives, and voiceless stops have longer constrictions than voiced stops (Slis & Cohen 1969).

3. *Acoustic duration of the noise burst of the stop*
Voiceless stops generally have longer noise bursts than voiced stops (Slis & Cohen 1969).

4. *Sound pressure of the noise burst and the friction noise*
The frequency components of a fricative and the burst of a stop have various origins. The components resulting from the rapid air flow through the vocal tract to the outside generally have higher intensities if the obstruent is voiceless than if it is voiced. The difference results from the fact that the glottis tends to be open during the realization of voiceless obstruents, whereas it is closed during voiced obstruents. When it is open, more air escapes simultaneously from the vocal tract with a higher velocity, which produces a louder sound when the air passes through the channel of the fricative, or the channel formed during the release of the constriction of the stop (Slis & Cohen 1969).

5. *Acoustic duration of the preceding vowel*
Voiceless obstruents are generally preceded by acoustically shorter vowels than voiced obstruents. According to Slis & Cohen (1969), this cue is not as important for the voiced/voiceless distinction in Dutch as it is in English.

6. *Sound pressure of the vowels adjoining the consonant*
Vowels adjacent to voiced obstruents have slightly higher amplitudes than vowels adjacent to voiceless obstruents (Lehiste & Peterson 1959). In Dutch, there is particular a relation between the voiced/voiceless distinction of an obstruent and the amplitude of the preceding vowel (Slis & Cohen 1969).

7. *Duration and spectral extensiveness of the vowel formant transitions*
The formant transitions from a vowel to a following stop, and from a stop to a following vowel tend to be longer and larger if the stop is voiced than if it is voiceless (Delattre 1962; Cooper et al. 1952; Slis & Cohen 1969). This cue is of minor importance for the voiced/voiceless distinction in Dutch.

8. *The peak value of the fundamental frequency in the following vowel*
   The fundamental frequency decreases after obstruents. The fall starts from a higher frequency and ends lower after a voiceless obstruent than after a voiced one (see for example Ohde 1984).

9. *Decay and rise time of the surrounding vowels*
   The intensity of a vowel declines more slowly before a voiced obstruent than before a voiceless one, and rises more slowly after a voiced one (Debrock 1977).

These acoustic cues are not equally important. In some contexts, some cues are more important than others. Van den Berg (1988) showed that the presence or absence of vocal fold vibration is the most important cue for obstruents in clusters. The initial obstruent of a cluster is perceived as voiced when glottal vibration is present during its last part, or during the first part of the following obstruent. The final obstruent is perceived as voiced when glottal vibration is present during its last part. It is not exactly known which cues are the most important for obstruents in other contexts.

There are trading relationships between the different cues (Slis & Cohen 1969; Rietveld & van Heuven 1997: 64). That is, if one (important) cue provides hardly any information on the voiced/voiceless distinction, listeners make the distinction on the basis of other cues. Speakers sometimes realize these other cues more clearly in order to compensate for the uselessness of the first one.

7.4 Coda and word-final obstruents

7.4.1 Introduction

Section 7.2 concluded that analyses of the realization of obstruents in Dutch as voiced or voiceless in Dutch should assume that the realization of final obstruents is not completely determined in phonology. In the remainder of this chapter, we will propose such an analysis.

Following Booij, we assume that all neutral obstruents behave alike, and should consequently be accounted for by the same process. The fact is that not only the word-final, but also the word-medial obstruents in coda positions tend to be voiced before voiced stops, and voiceless in most other contexts, and that all word-final obstruents — those in coda positions as well as those in onset positions — can be realized unfaithfully to their underlying [voice]-specifications (§3.4.2, §3.4.3, and §4.4.2).

The discussion of the analysis will start with the part that is relevant to neutral obstruents. First, we will give an informal formulation of the hypothesis which forms the basis of this part (§7.4.2), and show that it can account for all known facts (§7.4.3 and §7.4.4). We will then give a precise formulation of the hypothesis within the phonological theory adopted in this study (§7.4.5). Finally, we will mention some hypotheses which are implied by this part of the analysis (§7.4.6). The part of the analysis which is relevant to non-neutral obstruents will be discussed in section 7.5.

7.4.2 Two types of analysis

7.4.2.1 Introduction

Two types of analyses of the realization of obstruents in Dutch are possible which assign roles to both phonology and phonetics. The first one assumes that all obstruents are specified for [voice] in the phonological form, and that the realization of neutral obstruents is determined by these [voice]-specifications in combination with factors influencing phonetic implementation. This type of analysis will be discussed in section 7.4.2.2. The second type of analysis assumes that obstruents are specified neither as [+voice] nor as [-voice] in the output of phonology, and that their realization is completely determined by factors influencing phonetic implementation (§2.4.2). This type will be discussed in section 7.4.2.3.

7.4.2.2 Neutral obstruents have phonological specifications

Adherents of the assumption that neutral obstruents are specified for [voice] in the output of phonology have to decide on the phonological [voice]-specifications of these obstruents. The most obvious candidates are listed in (2).

(2) Candidates for the phonological [voice]-specifications of neutral obstruents

a. The underlying [voice]-specifications;
b. The [voice]-specification with which the obstruents surface in most contexts, i.e. [-voice];
c. The [voice]-specification which is universally less marked in most contexts, i.e. [-voice];
d. The [voice]-specification which is less marked in the relevant contexts. This specification is [+voice] in some contexts, including in between vowels, and [-voice] in others (as argued by Wetzels 1994, 1997 for Limburg Dutch and Bakairi respectively).

Choosing among these candidates is difficult, since they are all problematic in some respect. Analyses adopting candidates (2a) and (2d) cannot be correct for Standard Dutch since they are unable to explain why many speakers claim that both underlyingly voiced and underlyingly voiceless word-final obstruents are realized as voiceless when they follow vowels and precede vowel-initial words (§4.2.2). The assumption is that speakers' intuitions concern the lexical level (§4.2.1). If the candidate specification (2a) or (2d) is correct, neutral obstruents can be specified as [+voice] at this level, which predicts that speakers will assume that all neutral obstruents which are underlyingly voiced (in case of 2a), or all neutral obstruents in intervocalic positions (in case of 2d) can be realized as voiced.

Phonological specifications (2b-d) are problematic since analyses adopting them must assume that neutral obstruents are distinguished from the faithful obstruents first in phonology and then anew in phonetics. They first have to be distinguished in phonology, since only the neutral ones should be specified in accordance with (2b-d). Being specified as such, the neutral obstruents cannot be distinguished from the faithful obstruents in the output of phonology, since if specified as [-voice], they have the same [voice]-specification as all underlyingly voiceless faithful obstruents, and, if specified as [+voice], they have the same specifications as the underlyingly voiced faithful obstruents. The neutral obstruents have to be distinguished from the faithful obstruents in phonetics, since it is only the neutral ones that need not be realized in accordance with their phonological specifications.

Finally, analyses adopting phonological [voice]-specifications (2b) and (2c) assume that underlyingly voiced obstruents preceding voiced stops in compounds which are stored in the lexicon, such as *zandbank* /zɑnd+bɑŋk/ 'sandbank', are specified as [-voice] in the output of phonology, despite the fact that they are voiced in the underlying form, and nearly always voiced in the acoustic form. One may wonder whether this is unproblematic for language acquisition, and whether we can accept the implication that phonology first causes all neutral obstruents to be [-voice], whereas some of the underlyingly voiced ones are always realized as voiced.

In conclusion, the obvious phonological [voice]-specifications for neutral obstruents are all problematic. This implies that analyses assuming phonological specifications for these obstruents are problematic too. Apparently, we need a analysis which assumes that these obstruents have no phonological [voice]-specification at all.

7.4.2.3 The Complete Neutralization Hypothesis

The assumption that neutral obstruents do not have [voice]-specifications in the output of phonology and that their realization is determined by phonetics will be referred to as the Complete Neutralization Hypothesis (see 3).[2]

(3) **Complete Neutralization Hypothesis (CNH)**
Neutral obstruents do not have [voice]-specifications in the output of phonology, and their realization is completely determined by phonetics.

The CNH implies that neutral obstruents are voiced when the voiced realization requires no additional articulatory effort, and as voiceless when the voiceless realization requires no additional articulatory effort. The CNH, therefore, assumes that the realization of neutral obstruents is completely determined by the speaker's natural tendency to expend as little articulatory effort as possible.

Under the CNH, obstruents that are categorically realized as voiceless must be phonologically specified as [-voice]. They cannot be unspecified for [voice] because they would then be expected to be sometimes voiced and sometimes voiceless. This implies that the hypothesis is incompatible with the assumption that [voice] is a privative feature and the specification [-voice] does not exist, as is claimed by Mester & Itô (1989), Cho (1990), and Lombardi (1995a), among many others. Support for the assumption that [-voice] can be present in phonology has been provided by Lombardi (1996) and Wetzels (1994, 1997). They show that various processes in a number of languages cannot be described without reference to [-voice].

---

[2] In Ernestus (forthcoming) this hypothesis is referred to as the Permanent Neutralization Hypothesis.

Analyses of the realization of obstruents as voiced or voiceless in Dutch which adopt the CNH do not face the problems beset by analyses which assume that neutral obstruents are specified for [voice] in the output of phonology. First, they can explain the intuitions of many speakers that all word-final obstruents are generally realized as voiceless. Neutral obstruents are probably unspecified as early as at the lexical level of phonology (§7.4.5), and when speakers are forced to classify unspecified obstruents, they probably classify them with the default specification [-voice].

Second, within analyses assuming the CNH, the distinction made in phonology between neutral and faithful obstruents is accessible to phonetics. The former obstruents have no [voice]-specifications in the output of phonology, whereas the latter ones are specified as [+voice] or [-voice]. The two types are therefore clearly distinct in the input of phonetics, and the phonetic component does not need to mark obstruents as neutral or non-neutral all over again.

Finally, analyses adopting the CNH do not have to assume that obstruents in the output of phonology have a [voice]-specification which is the exact reverse of both the underlying one and the one with which the obstruent nearly always surfaces in the acoustic output.

In conclusion, analyses of [voice] in Dutch which assume the CNH are theoretically more adequate than other analyses. This implies that they should be adopted if they can account for at least the same data as other analyses. This will be seen to be the case in section 7.4.3, which argues that analyses adopting the CNH predict that neutral obstruents are generally realized as voiced before voiced stops, and as voiceless in all other contexts. Section 7.4.4 will show that these analyses also account for the influences of the factors mentioned in (1) in section 7.2 on the realization of word-final obstruents before voiced stops.

7.4.3 General realization of unspecified obstruents

7.4.3.1 Introduction

Adequate analyses of [voice] in Dutch must predict that neutral obstruents are generally realized as voiced before voiced stops, and as voiceless in all other contexts. This implies that analyses assuming the Complete Neutralization Hypothesis are adequate only if they predict that the phonetic component generally realizes unspecified obstruents as voiced before voiced stops, and as voiceless in all other contexts. Sections 7.4.3.2 to 7.4.3.5 will show that the phonetic component indeed realizes unspecified obstruents as voiced before voiced stops and as voiceless before voiceless obstruents, vowel-initial content words, and major phonological boundaries. The realization of unspecified obstruents before sonorants will be discussed in section 7.4.3.6.

Note that it is unknown what prediction analyses should make with respect to the realization of neutral obstruents before vowel-initial function words, since it is not known whether these obstruents are generally realized as voiced or voiceless (§3.4.3 and §4.2.2). Analyses do not have to make predictions with respect to the realization of neutral obstruents before voiced fricatives, since neutral obstruents never precede them in Dutch (§3.4.4).

7.4.3.2 Obstruents before voiced stops

Reduction in articulatory effort means (partial) gestural overlap (§1.1). As a consequence, unspecified obstruents are expected to be (partly) coarticulated with adjacent segments with respect to glottal vibration. Phonologically voiced stops must be realized as voiced, and when following unspecified obstruents, i.e. when being at the end of obstruent clusters, they are therefore realized with glottal vibration at least during their last parts (§7.3.4). This implies that when an unspecified obstruent precedes a phonologically voiced stop, it is coarticulated with a stop which is partly realized with glottal vibration. In the resulting realization, the vocal folds tend to vibrate well before the release of the voiced stop. The first phase of the voiced stop, or even the last phase of the preceding unspecified obstruent, is realized with glottal vibration, and the unspecified obstruent is perceived as voiced (§7.3.4). In other words, the phonetic component generally realizes unspecified obstruents as voiced before voiced stops, and analyses adopting the CNH predict the observed fact that neutral obstruents tend to be voiced before voiced stops.

7.4.3.3 Obstruents before voiceless obstruents

Phonologically voiceless obstruents at the end of obstruent clusters are realized without glottal vibration at least during their last parts (§7.3.4). Since unspecified obstruents are partly coarticulated with adjacent segments with respect to glottal vibration, a cluster of an unspecified obstruent and a phonologically voiceless obstruent tends to be realized without vocal fold vibration well before the last part of the phonologically voiceless obstruent. The largest part of the phonologically voiceless obstruent, or even the last phase of the preceding unspecified obstruent, is realized without glottal vibration, and the unspecified obstruent is perceived as voiceless (§7.3.4). In other words, the phonetic component generally realizes unspecified obstruents as voiceless before voiceless obstruents, and analyses adopting the CNH predict the observation that neutral obstruents before voiceless obstruents tend to be voiceless.

7.4.3.4 Obstruents before vowel-initial content words

Vowel-initial content words are generally preceded by glottal stops in the phonetic form (Jongenburger & van Heuven 1991). Therefore, obstruents preceding vowel-initial content words are part of obstruent clusters and the presence/absence of glottal vibration probably determines whether they are perceived as voiced or voiceless. For the realization of a glottal stop, the vocal folds have to be firmly pressed onto each other, which impedes glottal vibration. This means that unspecified obstruents, which are coarticulated with adjacent glottal stops with respect to glottal vibration, are at least partly realized with a more or less firmly closed glottis, that is, without glottal vibration, and are perceived as voiceless. In other words, the phonetic component realizes unspecified obstruents as voiceless before vowel-initial content words, and analyses adopting the CNH make the correct prediction with respect to the realization of such obstruents in these contexts.

If vowel-initial words are not realized with a glottal stop, analyses adopting the CNH still predict that preceding obstruents may be realized as voiceless. Possible causes then include the presence of preceding obstruents (cf. §7.4.4.2), and the length of the unspecified obstruent (see Hypothesis 5 in §7.4.6).

7.4.3.5 Obstruents before major phonological boundaries

Obstruents preceding major phonological boundaries, such as Intonational Phrases, are acoustically relatively long (Wightman et al. 1992; Cambier-Langeveld 1997, 2000), and long obstruents tend to be perceived as voiceless (§7.3.4). Obstruents preceding important phonological boundaries are therefore generally perceived as voiceless, unless special action is taken in order to make them sound as voiced. No such action is taken in the case of unspecified obstruents. The phonetic component consequently realizes unspecified obstruents before major phonological boundaries as voiceless. In other words, analyses adopting the CNH also make the correct prediction with respect to the realization of neutral obstruents preceding such boundaries.

7.4.3.6 Obstruents before sonorants

For analyses adopting the CNH to be valid, they should predict that the phonetic component realizes unspecified obstruents as voiceless before sonorants, since neutral obstruents tend to be voiceless before this type of consonant. So far, little is known of the acoustic cues to the voiced/voiceless distinction of obstruents in presonorantal positions in Dutch. It is therefore unknown whether the phonetic component realizes unspecified obstruents as voiced or as voiceless in these positions, and whether the analyses make the correct prediction.

7.4.3.7 Conclusions

In conclusion, there is no evidence that neutral obstruents in Dutch must be phonologically specified as [−voice] or [+voice] in order to be realized as voiced or voiceless in the contexts in which they are generally perceived as such. For all contexts for which the principal acoustic cues to the voiced/voiceless distinction are known, the realizations of these obstruents is also predicted if they are unspecified for [voice]. It seems, then, that analyses adopting the CNH can account for the realizations of neutral obstruents in Dutch.

7.4.4 Factors influencing the realization of unspecified obstruents

7.4.4.1 Introduction

Section 7.2 mentioned several factors that influence the realization of word-final obstruents before voiced stops: type of preceding segment, presence of stress, speech rate, speaker's sex, and speaker's mood. A valid analysis should be able to account for the influence of all of them. We will argue in sections 7.4.4.2 to 7.4.4.5 that such an analysis may be constructed if the Complete Neutralization Hypothesis is adopted.

7.4.4.2 Type of preceding segment

We saw in section 7.4.3.2 that analyses adopting the CNH predict that neutral obstruents before phonologically voiced stops are generally perceived as voiced. They are partly coarticulated with the following voiced stop, which is realized with glottal vibration at least during its last part. As a consequence of the coarticulation, the first part of the voiced stop may be realized with glottal vibration as well, in which case the unspecified obstruent is perceived as voiced.

The first part of the stop is most likely realized with glottal vibration if the entire obstruent cluster is realized with glottal vibration. If glottal vibration is absent for any length of time, it has to be started before the last part of the phonologically voiced stop, and it occasionally happens not to start well before the release of this stop.

Unspecified obstruents in coda positions are always preceded by a sonorant or are part of a cluster of unspecified obstruents preceded by a sonorant. Since sonorants are realized with a closed glottis, as are phonologically voiced stops in obstruent clusters, (clusters of) unspecified obstruents between such segments are realized with a closed glottis as well. This implies that all clusters consisting of an unspecified obstruent (cluster) in coda position and a voiced stop can be realized with continuous glottal vibration. The lengths of the unspecified obstruent, the phonologically voiced stop and the number of unspecified obstruents in the cluster determine the probability with which the cluster is realized with continuous glottal vibration, and therefore with which it is realized as completely voiced.

If the cluster consists of a single unspecified obstruent and a phonologically voiced stop which are both acoustically short, the cluster tends to be realized with ongoing vibration. If the unspecified obstruent happens to be a stop, the cluster is realized as a short constriction followed by the release of the phonologically voiced stop. The air pressure built up in the vocal tract during this constriction is too low to impede the opening of the vocal folds, and the stop cluster is realized with ongoing glottal vibration. If the obstruent happens to be an acoustically short fricative, the acoustic requirements for friction are met during such a short period that they do not prevent the vocal folds from vibrating. Since the short constriction of the phonologically voiced stop does not prevent glottal vibration either, obstruent clusters only consisting of a short unspecified fricative and a short phonologically voiced stop are realized with ongoing glottal vibration as well.

In contrast, clusters consisting of an acoustically long obstruent and a phonologically voiced stop are not realized with continuous glottal vibration. If the unspecified obstruent is a long stop, the cluster is not realized with ongoing glottal vibration since the air pressure built up in the vocal tract during the long constriction of the stop impedes glottal vibration. If the unspecified obstruent is a long fricative, glottal vibration is stopped because it is difficult to meet simultaneously the aerodynamic requirements for fiction and vibration for any extended period (§7.3.4).

In addition, clusters consisting of several unspecified obstruents and a phonologically voiced stop tend not to be realized with continuous glottal vibration. If the unspecified obstruent before the phonologically voiced stop is a stop and is preceded by another unspecified stop, it is part of an stop cluster. Such a cluster is realized with a long constriction, which impedes glottal vibration. If the unspecified obstruent before the voiced stop is a fricative preceded by another unspecified fricative, it is part of a fricative cluster. Such clusters are realized with a long period of friction, which cannot easily be realized with glottal vibration. Finally, if the unspecified obstruent before the voiced stop is a stop which follows an unspecified fricative, or a fricative which follows an unspecified stop, it is part of a cluster that is realized with a period of friction and a long constriction, or with two constrictions and a period of friction, respectively. Since both friction and constriction impede glottal vibration, such combinations probably constitute a serious obstruction for glottal vibration.

In summary, unspecified obstruents that are not acoustically long followed by phonologically voiced stops are generally realized with continuous glottal vibration. In contrast, long unspecified obstruents and clusters of unspecified obstruents in combination with a following voiced stop tend to be realized with a period without glottal vibration. If glottal vibration is absent, it may occasionally not start well before the release of the phonologically voiced stop and the unspecified obstruents may be perceived as voiceless. This is why the phonetic component realizes non-long unspecified obstruents invariably as voiced before voiced stops, whereas long unspecified obstruents and clusters of unspecified obstruents are realized as voiced in these positions most of the time, but not invariably.

In conclusion, analyses adopting the CNH predict that neutral obstruents which are not acoustically long and precede voiced stops are nearly always voiced if they follow voiced segments. If they follow other, neutral, obstruents, they are sometimes realized as voiceless. This is in conformity with the data (§7.2).

7.4.4.3 Presence of stress

Faithfulness of the phonological output to the input is generally more important for stressed syllables than for unstressed ones (cf. Beckman 1997). It is probable that faithfulness not only plays a role in phonology, but also in phonetics. In that case, segments of stressed syllables are more often realized faithfully to their phonological specifications than segments of unstressed syllables, which implies that a phonologically voiced stop in an obstruent cluster is more often realized with glottal vibration at full strength, and glottal vibration is more often present during a large part of the realization of a cluster consisting of an unspecified obstruent and a voiced stop, if the stop is part of a stressed syllable. This in turn would mean that the phonetic component is more likely to realize obstruent clusters consisting of an unspecified obstruent and a voiced stop as voiced when the stop is part of a stressed syllable than when it is part of an unstressed one. Hence, analyses adopting the CNH make the correct prediction with respect to the influence of stress (§7.2).

There is an additional reason why the phonetic component probably tends to realize obstruent clusters ending in stressed onset obstruents as voiced. Stress involves a higher subglottal pressure in about the first half of the syllable, at least in English (Patricia Keating, personal communication). This high pressure favours the presence of vocal fold vibration, and therefore the perception of obstruent clusters as voiced. This means that analyses adopting the CNH predict the influence of stress also on purely phonetic grounds.

7.4.4.4 Speech rate

When the speech rate increases, segments become shorter. When a stop is shorter, its constriction is shorter, and the air pressure built up in the vocal tract during its constriction is less likely to impede glottal vibration. When a fricative is shorter, the aerodynamic requirements for friction are less likely to prevent the vocals folds from vibrating (§7.3.4). Therefore, the phonetic component realizes unspecified obstruents before voiced stops more often with continuous glottal vibration, that is as voiced, when the speech rate is high than when the speech rate is low. Consequently, analyses adopting the CNH correctly predict that unspecified obstruents before voiced stops are more likely to be voiced in fast speech.

If an unspecified obstruent nevertheless happens to be realized without glottal vibration, less time is available for the vocal folds to start vibrating again before the release of the following voiced stop in fast speech than in slow speech. When the vocal folds do not vibrate in time, and start to vibrate only after the release of the stop, the entire obstruent cluster is perceived as voiceless. For this reason, the phonetic component realizes obstruent clusters consisting of an unspecified obstruent and a voiced stop as completely voiceless more often in fast speech than in slow speech.

This is particularly true in the case of clusters starting with fricatives. Friction generally prevents the vocal folds from vibrating, whereas the complete constrictions of the stops only do so when they are long (§7.3.4). As a consequence, particularly in clusters starting with fricatives, the vocal folds occasionally do not vibrate during the release of the phonologically voiced stop, which is why it is particularly these clusters that are occasionally realized by the phonetic component as completely voiceless. We find that analyses adopting the CNH predict that particularly clusters of unspecified fricatives and voiced stops are more likely to be realized as completely voiceless in fast speech than in slow speech, which is again in line with the data (§7.2).

7.4.4.5 Speaker's sex

Men have longer and heavier vocal folds than women. According to Slis (1987), this implies that it is easier for men than for women to realize obstruent clusters with ongoing vocal fold vibration. In addition, men generally differ from women in that their vocal tract is larger, which means that more air can flow from their lungs into the vocal tract before the difference between the subglottal and supraglottal pressure is too low for the subglottal pressure to open the glottis, and glottal vibration stops (Vincent van Heuven, personal communication). In other words, the phonetic component of a male speaker is more likely to realize clusters of unspecified obstruents and voiced stops with continuous glottal vibration than the phonetic component of a female speaker. This means that analyses incorporating the CNH make the correct prediction that men more often realize neutral obstruents before voiced stops as voiced (§7.2).

7.4.4.6 Speaker's mood

When speakers are in an emotional state, more especially when they are angry, they tend to speak with force (Meinsma 1958; Demeulemeester 1962), i.e. with a strong air stream from the lungs into the vocal tract. This strong air stream prevents the glottis from closing. Emotional speakers therefore tend to realize unspecified obstruents without glottal vibration. As a consequence, their unspecified obstruents in obstruent clusters tend to sound as voiceless. This implies that analyses adopting the CNH predict an influence of the speaker's mood which is in line with the data (§7.2).

7.4.5 The analysis

The above sections discussed the Complete Neutralization Hypothesis, and showed that it can account for all observed Dutch data. The hypothesis was informally formulated in 7.4.2.3. We will now explicitly propose an analysis for the neutral obstruents in Dutch which adopts the hypothesis.

Analyses adopting the CNH must explicitly state which obstruents are unspecified for [voice]. It was argued in section 7.4.1 that all syllable-final and word-final obstruents, i.e. all neutral obstruents, behave alike. These obstruents differ from other obstruents in that they are in coda position at the lexical level. The facts can therefore be captured by an analysis which states that all obstruents which are in coda position at the lexical level are unspecified for [voice] in the output of the lexical and post-lexical level of phonology and in the input of phonetics. All other obstruents, with the exception of those which will be discussed in section 7.5, enter phonetics with their underlying [voice]-specifications. This analysis is illustrated in (4), in which "T" stands for an "alveolar stop unspecified for [voice]".

(4) *Illustration of the analysis*

Input lexical level of phonology:	$t)_\sigma$	$d)_\sigma$	$_\sigma(t$	$_\sigma(d$
Output lexical level of phonology:	T	T	t	d
Output post-lexical level of phonology:	T	T	t	d
Output phonetics:	[t] or [d]	[t] or [d]	[t]	[d]

This analysis can be incorporated into Optimality Theory as follows. Coda obstruents are assumed to be unspecified at the lexical level as the result of the high ranking of a constraint which forbids coda obstruents to be specified for [voice]. This constraint will be referred to as *CODA(VOICE) (see 5). Constraints holding only for coda obstruents have been proposed before by e.g. Itô (1989), and Mascaró & Wetzels (1999). We do not adopt a constraint which requires all obstruents, instead of only coda obstruents, to be unspecified for [voice] (contra Ernestus forthcoming), because this stronger constraint would not allow for the account of the regular past-tense morpheme which will be proposed in section 7.5.4.

(5) *CODA(VOICE): No coda obstruent has a [voice]-feature.

Not all obstruents are unspecified for [voice] in the output of phonology. Therefore, *CODA(VOICE) should directly dominate a constraint that requires all [voice]-features in the input to have a correspondent in the output, or all obstruents to be specified for [voice]. Possible candidates are the three constraints defined in (6).

(6) Constraint candidates

a. MAXFEATURE(VOICE): Every [voice]-feature in the input has a correspondent in the output.

b. IDENTITYFEATURE(VOICE): Corresponding segments are identical in [voice].

c. SPECFEATURE(VOICE): Every obstruent is specified for [voice].

The first constraint candidate is MAXFEATURE(VOICE), which states that [voice]-features in the input must be present in the output, but do not necessarily have the same value in the input and the output. Tableau 7.1 shows that the interaction of *CODA(VOICE) with MF(VOICE), i.e. constraint ranking (7), ensures that coda obstruents are unspecified for [voice] in the output of phonology, whereas onset obstruents are faithful to their underlying [voice]-specifications. In the tableau, "T" stands for an alveolar, and "P" for a bilabial stop which is unspecified for [voice].

(7) Constraint ranking at the lexical level:
*CODA(VOICE) >> MF(VOICE)

Tableau 7.1 Evaluation of output candidates of phonology: obstruents in onset and coda positions.

bad /bɑd/ 'bath'	*CODA(VOICE)	MF(VOICE)
(bɑd)$_\sigma$	*!	
(bɑt)$_\sigma$	*!	
☞ (bɑT)$_\sigma$		*
(PɑT)$_\sigma$		**!

Apparently, the constraint ranking in (7) represents well the CNH, and MAXFEATURE(VOICE) can be adopted for the analysis.

The second candidate is the constraint IDENTITYFEATURE(VOICE), which was defined by McCarthy & Prince (1995: 264). It states that corresponding segments are identical in the feature [voice] (6b). Because there is a difference between a [+voice] or [-voice]-specification and no specification, the constraint can be interpreted as preventing segments which are specified in the input from being unspecified in the output. Under this interpretation, it has at least the same effects as MAXFEATURE(VOICE), and in combination with *CODA(VOICE) is a good expression of the CNH. The interpretation is, however, not desirable. It implies that the input-output correspondence between a [+voice] and a [-voice] specification is always as bad as the input-output correspondence between a [+voice] or [-voice]-specification and no specification: both correspondences imply one violation of IDENTITYFEATURE(VOICE). This implication is probably incorrect, since there are many languages in which some segments which are [+voice] in the input can be [-voice] in the output, or vice versa, but cannot be unspecified. Apparently, the segments in these languages must satisfy a constraint which states that segments which are specified in the input are also specified in the output, whereas they are allowed to violate a constraint which states that corresponding [voice]-specifications in the input and output have the same value. That is, the constraint on the absence of features appears to differ from the constraint on the identity of corresponding feature specifications. Consequently, the constraint INDENTITYFEATURE(VOICE) is preferably not interpreted as also banning unspecified obstruents from the output which are specified in the input, and should not express the CNH with *CODA(VOICE).

The final constraint candidate for expressing the CNH with *CODA(VOICE) is SPECFEATURE(VOICE), which states that obstruents must be specified for [voice] (6c). It has exactly the same effects as MAXFEATURE(VOICE) for obstruents which are specified for [voice] in the input. Consequently, it expresses the CNH as well as MAXFEATURE(VOICE) in combination with *CODA(VOICE). Since we do not know of any data or theoretical grounds which favour SPECFEATURE(VOICE) above MAXFEATURE(VOICE), or vice versa, we arbitrarily opt for one of them, viz. MAXFEATURE(VOICE).

In conclusion, we propose an analysis for the realization of neutral obstruents in Dutch which assumes that obstruents which are in coda positions at the lexical level are unspecified for [voice] in the outputs of the lexical and post-lexical level of phonology. They are unspecified as a result of constraint ranking (7).

7.4.6 Hypotheses

The analysis for [voice] in Dutch proposed so far implies the following hypotheses which can be verified on the basis of speech data.

1. *No influence of underlying [voice]-specifications*
   The acoustic properties of neutral obstruents are not influenced by their underlying [voice]-specifications. Jongman et al. (1992) and Baumann (1995) have shown this to be true for utterance-final obstruents, and for word-final obstruents preceding nasals or the word *en* /ɛn/ 'and'. It should also be true for neutral obstruents in other contexts.

2. *No influence of phonological features*
   The phonological features of a neutral obstruent, or those of the segments adjacent to a neutral obstruent, can influence the realization of the neutral obstruent as voiced or voiceless, only if they correspond to articulatory gestures which happen to influence acoustic characteristics related to the perception of voicing. The phonological features cannot influence the realization of the

neutral obstruent as voiced or voiceless by phonological processes, since neutral obstruents have no phonological [voice]-specifications to begin with.

3. *No gestures*
No articulatory gestures are made in order to realize a neutral obstruent as voiced or voiceless. The only gestures being made are those necessary for the realization of the obstruent according to its phonologically specified manner and place of articulation, and those necessary for the realization of the following segment, if present.

4. *Differences in realization between neutral and non-neutral obstruents*
Neutral obstruents are realized as voiced more often than phonologically voiceless obstruents, and less often than phonologically voiced obstruents. Neutral obstruents are unspecified for [voice], and therefore can be realized as voiced and voiceless. In contrast, obstruents with phonological [voice]-specifications must be realized in accordance with these specifications.

5. *Word-final intervocalic obstruents*
Word-final obstruents which are post-vocalic and directly followed by a vowel in the phonetic form tend to be realized as voiced in fast speech, and as voiceless in slow speech. Segments realized in fast speech tend to be short, and those realized in slow speech long. In English, all the relatively short intervocalic obstruents are perceived as voiced, whereas the long ones tend to be perceived as voiceless (Lisker 1957). If the perception of intervocalic obstruents as voiced or voiceless is determined by length also in Dutch, which is a reasonable assumption, this means that intervocalic obstruents in fast speech tend to be perceived as voiced, with the exception of those which must be voiceless on phonological grounds. In contrast, most intervocalic obstruents in slow speech are perceived as voiceless, with the exception of those which must be voiced on phonological grounds.

6. *Influence of the lexicon*
There is a systematic difference in realization between word-final obstruents which precede vowel-initial clitics in word-combinations which may be retrieved as single units from the lexicon, and in word-combinations which are always computed from their parts.

Clitics are function words, such as *er* /ər/ 'there', which do not form prosodic words of their own but are incorporated into the preceding prosodic word (§2.3.4 and §3.6). When they are vowel-initial and are preceded by consonants, they form syllables with these consonants, and the consonants are in onset position.

If a vowel-initial clitic forms a syllable with a preceding obstruent at the lexical level, this obstruent is in onset position at the lexical level, and specified for [voice] in the output of phonology. The obstruent is realized faithfully to its underlying [voice]-specification. A vowel-initial clitic forms a syllable with the preceding obstruents at the lexical level only if it is part of a word-combination that is retrieved as a single unit from the lexicon. If, on the other hand, it is part of a combination that is computed from its parts, it cannot form a syllable with the preceding obstruents at the lexical level, as the lexical level then deals with every grammatical word in isolation. In that case, the preceding obstruents are in coda position in the output of the lexical level, and unspecified for [voice]. Their realization as voiced or voiceless is determined by which realization requires no additional articulatory effort. Hence, there is a systematic difference in phonological [voice]-specification, and therefore also in realization, between word-final obstruents before vowel-initial clitics which are part of word-combinations that are stored in the lexicon, and word-combinations which are always computed from their parts.

The prediction about the nature of the systematic difference in realization depends on which view is adopted on the lexical form of strings (§2.2.2). One view holds that the lexical form of a string contains the same segments as the lexical forms of its parts (Booij 1985). This view (referred to as View I in what follows) predicts that a word-final obstruent which is voiced in the underlying form of the separate word is also voiced in the lexical forms of strings containing this word. If this obstruent precedes a vowel-initial clitic in a string, it is therefore systematically realized as voiced when this string is retrieved as a single unit from the lexicon. Hence, the obstruent is realized as voiced before vowel-initial clitics more often in word-combinations which are often retrieved as single units from the lexicon than in other word-combinations. In contrast, word-final obstruents which are voiceless in the underlying forms of the separate words are less often realized as voiced before vowel-initial clitics in word-combinations which are often retrieved as single units from the lexicon. These predictions are summarized in Table 7.1.

Another view, to which we will refer as View II, holds that word-combinations have lexical forms representing the acoustic form with the highest frequency of occurrence (Bybee 1995, 1996; Bybee & Scheibman 1999). We saw above that obstruents unspecified for [voice] are often realized as voiced in fast, i.e. normal, speech if they are intervocalic (see Hypothesis 5). View II predicts, then, that post-vocalic word-final obstruents which precede vowels are [+voice] in the lexical form of word-combinations, and are realized as voiced. In addition, we saw above that obstruents unspecified for [voice] are at least partly realized without glottal vibration, i.e. as voiceless, if they follow other unspecified obstruents (§7.4.4.2). View II implies that word-final obstruents between obstruents and vowel-initial clitics are predicted to be specified as [−voice] in the lexical form of word-combinations, and to be realized as voiceless. These predictions are different from the predictions made by View I. They are also summarized in Table 7.1.

Table 7.1. Predictions of View I and View II with respect to the realization of word-final obstruents before vowel-initial clitics in stored and non-stored word-combinations.

	Realization of a word-final obstruent before a vowel-initial clitic in a	
	stored combination	non-stored combination
View I	•voiced, if the obstruent is voiced in the underlying form of the separate word; •voiceless, if the obstruent is voiceless in the underlying form of the separate word.	voiced or voiceless, depending on which realization requires no additional articulatory effort.
View II	•voiced, if the obstruent follows a sonorant; •voiceless, if the obstruent follows another obstruent.	same as under View I.

7.4.7 Conclusions

Adequate analyses of the realization of syllable- and word-final obstruents as voiced or voiceless in Dutch must assign an important role to phonetics. We propose the Complete Neutralization Hypothesis, which assigns a major part to the speaker's natural tendency to expend as little articulatory effort as possible. It states that all coda obstruents and word-final obstruents are phonologically unspecified for [voice], and that they are realized as voiced when a voiced realization requires no articulatory effort, and as voiceless when the voiceless realization requires no articulatory effort. In combination with the phonetic component, the CNH is able to account for at least the available data. We formulated an explicit analysis for the neutral obstruents by incorporating the hypothesis within Optimality Theory.

7.5 Syllable-initial obstruents

7.5.1 Introduction

An analysis for the realization of neutral obstruents should only be adopted if it allows for an adequate account of the realization of all obstruents as voiced or voiceless, including the morpheme-initial ones. The realization of the underlyingly voiced fricatives in onset positions, word-initial /d/s, and the initial stop of the regular past-tense morpheme may seem problematic for the analysis proposed in section 7.4. These obstruents are generally assumed to be realized as voiceless because the preceding coda obstruents are specified as voiceless. Under the CNH, these preceding coda obstruents are unspecified for [voice], and the voiceless realization of the following initial obstruents must have a different source.

Sections 7.5.2 to 7.5.4 present a new description of these cases, which is compatible with the CNH. Unless indicated otherwise, the term "obstruents" in these sections refers to word-initial or word-medial obstruents that are in onset positions.

7.5.2 An analysis for fricatives in onset positions

7.5.2.1 Single fricatives

Both underlyingly voiced and voiceless fricatives are realized as voiceless after obstruents, and according to their underlying [voice]-specifications in all other segmental contexts (cf. §3.4.4). Developing an analysis for their realizations requires, in the first place, determining whether these realizations are phonological or phonetic in nature.

One might argue that the realization of fricatives is completely determined in phonetics. The fact is that all fricatives are realized as voiceless after obstruents, and it is easier to realize them as voiceless than as voiced in these positions. Clusters of a fricative following another fricative are generally realized with a relatively long period of friction, and it is difficult to meet the nearly opposite aerodynamic requirements for friction and glottal vibration for a long time. Stop-fricative clusters contain a period of constriction and a period of friction, both of which do not favour glottal vibration. Hence, glottal vibration is often absent in obstruent clusters ending in an unspecified fricative, and the fricative tends to be perceived as voiceless.

Nevertheless, the realization of onset fricatives cannot be due to lack of phonological specification, since they are systematically realized as either voiced or voiceless. That is, there is little free variation in their realization in a certain context, which is not what would be expected if their realization was completely determined by phonetics. Moreover, underlyingly voiced fricatives are generally realized as voiced in utterance-initial positions, although it is not actually easier to realize them as voiced than as voiceless in these positions. Utterance-initial fricatives, then, cannot be voiced for phonetic reasons only. They are voiced in order to satisfy phonological constraints, and onset fricatives are phonologically specified for [voice].

The fact that underlyingly voiced fricatives are realized as voiceless after obstruents, whereas underlyingly voiced stops are not, may be regarded as another result of the difference between stops and fricatives with respect to the strength of their voiced/voiceless opposition (cf. §3.3). This view is adopted here, since it explains why exactly fricatives are realized unfaithfully to their underlying [voice]-specifications after obstruents: fricatives have the weakest voiced/voiceless opposition. This view suggests the following analysis within Optimality Theory.

The difference in strength of the voiced/voiceless opposition implies that the faithfulness constraint on the [voice]-specifications of fricatives is ranked lower than the faithfulness constraint on the [voice]-specifications of stops. We assume the two constraints in (8), and ranking (9).

(8) Relevant IDENT(VOICE) constraints:

a.	IDENTFRIC(VOICE):	If a fricative has a specification for [voice] in the output, this specification is identical to the one in the input. Abbreviated as IF(VOICE).
b.	IDENTSTOP(VOICE):	If a stop has a specification for [voice] in the output, this specification is identical to the one in the input. Abbreviated as IS(VOICE).

(9) Constraint ranking: IDENTSTOP(VOICE) >> IDENTFRIC(VOICE)

It may be assumed that IDENT(VOICE) is a family of constraints, and that each class of segments has to satisfy a different member. The ranking of a member is generally determined by the effort needed to realize and perceive the voiced/voiceless distinction on the relevant type of segment. If it takes little effort to express and perceive the distinction, the member ranks high; if keeping up the distinction requires a lot of effort, the relevant member of the IDENT(VOICE) family ranks low. Since the cues for the voiced/voiceless distinction on a type of segment are language specific, the effort needed to keep up the distinction on the different segment types is different for each language, and the ranking of the members of IDENT(VOICE) is language specific. In Dutch, IDENTSTOP(VOICE) dominates IDENTFRIC(VOICE) probably because glottal vibration is an important cue for voicing, and glottal vibration is more easily realized in stops than in fricatives (§7.3.4).

Since in Dutch underlyingly voiced stops are realized as voiced, i.e. faithfully to their underlying specifications, whereas underlyingly voiced fricatives are unfaithfully realized as voiceless after obstruents, the constraints IDENTSTOP(VOICE) and IDENTFRIC(VOICE) are probably separated by a constraint in this language which requires obstruents to be voiceless when following other obstruents. We assume constraint (10).

(10) NO VOICED OBSTRUENTS IN CLUSTERS (NVOC): an obstruent in a cluster is not voiced.

This constraint is plausible, as it is phonetically grounded: it is generally more difficult to realize obstruents in clusters as voiced than as voiceless, since these clusters are realized with long closures, long periods of friction, or with both a constriction and a friction, which impede glottal vibration (see also above, and §7.4.4.2). Moreover, the constraint is plausible because it is satisfied by nearly all obstruents in word-medial clusters in Dutch (Zonneveld 1983).

Since the onset fricatives are phonologically specified for [voice], they satisfy MAXFEATURE(VOICE) (see 5b). This implies that violations of NVOC are not avoided by means of violations of MAXFEATURE(VOICE), and MAXFEATURE(VOICE) must dominate NVOC.

The following constraint ranking appears to be in force.

(11) Constraint ranking at the post-lexical level:
MF(VOICE), IDENTSTOP(VOICE) >> NVOC >> IDENTFRIC(VOICE)

This ranking holds at least for the post-lexical level, since the information which word-initial fricatives are preceded by obstruents, and consequently which voiced fricatives violate NVOC, is only available at that level. There is no reason why we should not assume that the ranking also holds for the lexical level.

Under the CNH, neutral obstruents are unspecified for [voice] in the input of the post-lexical level of phonology, and should remain unspecified up to the phonetic component. A constraint is therefore needed at the post-lexical level which requires obstruents which are unspecified in the input to be unspecified in the output. This constraint can be DEPIO(VOICE), which is defined in (12). It is based on the constraint DEPIO defined by McCarthy and Prince (1995).

(12) DEPIO(VOICE): Every [voice]-feature in the output has a correspondent in the input.

The constraint is probably ranked as high as MF(VOICE) and IDENTSTOP(VOICE), as it is never violated.

The adopted constraint hierarchy at the post-lexical level is therefore:

(13) Constraint ranking at the post-lexical level:
DEPIO(VOICE), MF(VOICE), IDENTSTOP(VOICE) >> NVOC >> IDENTFRIC(VOICE)

Tableaux 7.2 to 7.4 illustrate evaluations at the post-lexical level. They show that constraint ranking (13) provides the correct outputs for stops and fricatives in onset and coda positions. The letters "S", "F", and "P" represent different types of obstruents unspecified for [voice]. Tableau 7.2 shows that fricatives in utterance-initial positions have the same [voice]-specifications in the optimal output candidate as in the input.

Tableau 7.2 Evaluation of output candidates of post-lexical phonology: An underlyingly voiced fricative in utterance-initial position.

zee /(ze)$_\sigma$/ 'sea'	DEPIO(VOICE)	MF(VOICE) IS(VOICE)	NVOC	IF(VOICE)
☞ (ze)$_\sigma$				
(se)$_\sigma$				*!
(Se)$_\sigma$		*!		

Tableau 7.3 shows that underlyingly voiced fricatives preceded by obstruents are voiceless in the optimal output candidate of the post-lexical level of phonology. The coda obstruent of the input is unspecified for [voice], because of the constraint ranking (7), which holds at the lexical level.

Tableau 7.3 Evaluation of output candidates of post-lexical phonology: An underlyingly voiced fricative after an obstruent.

stoepzout /(stuP)$_\sigma$(zɔuT)$_\sigma$/ 'pavement salt'	DEPIO(VOICE)	MF(VOICE) IS(VOICE)	NVOC	IF(VOICE)
(stuP)$_\sigma$(zɔuT)$_\sigma$			*!	
☞ (stuP)$_\sigma$(sɔuT)$_\sigma$				*
(stup)$_\sigma$(sɔuT)$_\sigma$	*!			*
(stuP)$_\sigma$(SɔuT)$_\sigma$		*!		

Finally, Tableau 7.4 shows that constraint ranking (13) ensures that stops in onset positions have the same [voice]-specifications in the input and the output. The coda obstruent in the input is unspecified for [voice], because of ranking (7).

Tableau 7.4 Evaluation of output candidates of post-lexical phonology: An underlyingly voiced stop after an obstruent.

afbeelden /(ɑF)$_\sigma$(bel)$_\sigma$(dən)$_\sigma$/ 'to depict'	DEPIO(VOICE)	MF(VOICE) IS(VOICE)	NVOC	IF(VOICE)
☞ (ɑF)$_\sigma$(bel)$_\sigma$(dən)$_\sigma$			*	
(ɑF)$_\sigma$(pel)$_\sigma$(dən)$_\sigma$		*!		
(ɑf)$_\sigma$(pel)$_\sigma$(dən)$_\sigma$	*!	*		
(ɑF)$_\sigma$(Pel)$_\sigma$(dən)$_\sigma$		*!		

7.5.2.2 Fricative geminates

The overall analysis proposed above can account for the realization of fricatives in all contexts, except before or after a fricative with the same place of articulation. Clusters of fricatives of the same place of articulation arise when a word-final fricative is followed by a word-initial one. These clusters are generally realized with a duration that is shorter than the duration of two segments (§3.5). If the analysis has to account for the duration of these clusters, it needs to be extended. In what follows, clusters consisting of two segments with the same manner and place of articulation will be referred to as geminates.

The realization of fricative geminates as single obstruents cannot be adequately explained with a phonological constraint NOGEM, which forbids geminates. The problem is that fricative geminates are always realized as voiceless, independently of their context, exact duration, etc. They therefore must be specified as [-voice] in the output of phonology. If NOGEM is a phonological constraint, they are not always phonologically voiceless. Phonological NOGEM must be in force at the post-lexical level, since word-final segments are not followed by word-initial ones before this level. Whatever the position of NOGEM in the constraint hierarchy at the post-lexical level of phonology, an input containing an underlyingly voiced or voiceless fricative followed by a voiced one has, incorrectly, a voiced fricative as its output. This is shown in Tableaux 7.5 and 7.6. The left-pointing finger in these tableaux indicates the optimal output of the relevant constraint ranking, whereas the right-pointing finger indicates the actual output. The letter "S" denotes an alveolar fricative unspecified for [voice]. The first fricative is in coda position, and is therefore unspecified for [voice] in the input of the post-lexical level. As a consequence, the fricative cluster has only the [voice]-specification of the second fricative in the input, and IF(voice) ensures that the corresponding segment in the output has this same [voice]-specification.

Tableau 7.5 Evaluation of output candidates of post-lexical phonology: A fricative geminate. NoGem *is high in the constraint hierarchy.*

los zijn /lɔS zɛin/ 'loose are'	NoGem DepIO(voice)	MF(voice) IS(voice)	NVOC	IF(voice)
lɔSsɛin		*!		*
☜ lɔzɛin				
lɔSɛin		*!		
☞ lɔsɛin				*!

Tableau 7.6 Evaluation of output candidates of post-lexical phonology: A fricative geminate. NoGem *is low in the constraint hierarchy.*

los zijn /lɔS zɛin/ 'loose are'	DepIO(voice)	MF(voice) IS(voice)	NVOC	IF(voice)	NoGem
lɔSsɛin				*	*
☜ lɔzɛin					
lɔSɛin		*!			
☞ lɔsɛin				*	

Apparently, under the assumption that NoGem ensures the correct length of fricative geminates in phonology, an ad hoc constraint is necessary in order to account for the phonological [−voice]-specification of these geminates.

In contrast, the realization of fricative geminates as single (long) segments can be adequately accounted for, without ad hoc assumptions, in phonetics. The fact is that if the degemination process is phonetic, it has as its input the output of post-lexical phonology, i.e. the output of constraint ranking (13). In the output of post-lexical phonology, the fricative in coda position is unspecified for [voice] (§7.4.5), while the following fricative in onset position is [−voice] (§7.5.2.1). This is also the case if the input of this level consists of a word-final fricative and an initial underlyingly voiced one, as is illustrated in Tableau 7.7, in which "F" denotes an unspecified labiodental fricative. Since the input to the phonetic process of degemination is always a cluster consisting of an unspecified fricative and a voiceless one, the output is always a single voiceless fricative. The fricative is voiceless, as degemination does not influence the realization of the feature specifications of the fricative cluster. These feature specifications are realized faithfully.

Tableau 7.7 Evaluation of output candidates of post-lexical phonology: A fricative geminate. NOGEM is absent.

afvegen /$(\text{ɑF})_\sigma(\text{ve})_\sigma(\text{xən})_\sigma$/ 'to wipe'	DEPIO(VOICE)	MF(VOICE), IS(VOICE)	NVOC	IF(VOICE)
$(\text{ɑF})_\sigma(\text{ve})_\sigma(\text{xən})_\sigma$			*!	
☞ $(\text{ɑF})_\sigma(\text{fe})_\sigma(\text{xən})_\sigma$				*
$(\text{ɑf})_\sigma(\text{fe})_\sigma(\text{xən})_\sigma$	*!			*
$(\text{ɑF})_\sigma(\text{Fe})_\sigma(\text{xən})_\sigma$		*!		

The assumption that degemination is phonetic is not ad hoc itself, since Martens & Quené (1994) have demonstrated that degemination is gradual, i.e. assigns durations to obstruent clusters between the length of one and two single obstruents, and so exhibits an important characteristic of phonetic processes.

In conclusion, the assumption that the degemination process is phonetic in nature can adequately account for the realization of fricative geminates.

7.5.2.3 Conclusions

We propose an analysis of the realization of onset fricatives which is founded on the claims that

- the realization of onset fricatives as voiced or voiceless is determined in phonology;
- fricatives can be realized unfaithfully to their underlyingly [voice]-specifications because their voiced/voiceless opposition is weak;
- obstruents in clusters are preferably not voiced;
- degemination is phonetic in nature.

The second claim explains why only fricatives are realized unfaithfully to their underlying [voice]-specifications, a fact unaccounted for by previous analyses. The proposed analysis does not refer to the [voice]-specifications of coda obstruents, and is therefore compatible with the CNH.

7.5.3 An analysis for word-initial /d/

The initial /d/ of several function words is the second type of initial obstruent of which the realization has been claimed to depend on the [voice]-specification of the preceding coda obstruent This /d/ is sometimes realized as voiceless after obstruents (§3.4.5).

The realization of /d/ is not determined by phonetics, as is evident from the fact that this segment is systematically realized as voiced in utterance-initial positions, whereas utterance-initial obstruents are not more easily realized as voiced than as voiceless.

Nevertheless, the realization of initial /d/ as [t] seems to be a case of hypo-articulation. This is suggested by the fact that it is only the /d/ of function words that may be realized as voiceless after obstruents. Function words are rarely crucial to the propositional content of the utterance, and are generally highly frequent. The speaker may therefore hypo-articulate them to a greater extent than content words without running the risk of being misunderstood.

The view adopted here, then, is that the voiceless realization of /d/ is a case of hypo-articulation. This view is compatible with the conclusion that the realization of /d/ is not determined by phonetics, if the hypo-articulation is assumed to be phonologized. This assumption implies that the voiceless realization of /d/ after obstruents is due to a phonological wellformedness constraint which results in a decrease in articulatory effort.

We assume that the relevant /d/ is phonologically voiceless after obstruents because it has to satisfy a member of the constraint family IDENT(VOICE) which is ranked lower in the hierarchy than NVOC. Recall from section 7.5.2.1 that NVOC forbids voiced obstruents in obstruent clusters (see 10). NVOC has as its effect a decrease in articulatory effort, since it is more difficult to realize obstruents as voiced than as voiceless in obstruent clusters (§7.4.4.2). The member of IDENT(VOICE) on the relevant /d/s will be referred to as IDENTHYPOSTOP(VOICE) (see 14).

(14) IDENTHYPOSTOP(VOICE): If a stop of a word that can be hypo-articulated to some extent has a specification for [voice] in the output as well as in the input, these two specifications are identical. Abbreviated as IHS(voice).

Since segments which can be hypo-articulated to some extent are less likely to be realized faithfully than other segments, the IDENT(VOICE) on these segments must be ranked lower than the IDENT(VOICE) on segments which cannot be hypo-articulated. Hence, IdentHypoStop(voice) is universally dominated by IDENTSTOP(VOICE), which was defined in (8b).

(15) Universal ranking: IDENTSTOP(VOICE) >> IDENTHYPOSTOP(VOICE)

The words of which the stops must satisfy IDENTHYPOSTOP(VOICE), instead of IDENTSTOP(VOICE), are marked in the input of post-lexical phonology. Their marks are probably the result of two facts: they contribute little to the propositional content of the utterance, so that the listener can do without them, and secondly, they are highly frequent, which ensures easy recognition. These two facts allow the speaker to hypo-articulate them to some extent without running the risk of being misunderstood.

The constraint NVOC dominates IDENTHYPOSTOP(VOICE) at least at the post-lexical level, since at this level words are not dealt with in isolation, so that it can be determined whether a word-initial /d/ violates NVOC. Given the constraint ranking arrived at in section 7.5.2.1, the resulting constraint ranking for this level is (16).

(16) Constraint ranking at the post-lexical level:
DEPIO(VOICE), MF(VOICE), IDENTSTOP(VOICE) >> NVOC >> IDENTFRIC(VOICE), IDENTHYPOSTOP(VOICE)

Tableau 7.8 shows that this constraint ranking can account for the phonological [-voice]-specification of a post-obstruental initial /d/ which is part of a word that can be hypo-articulated to some extent. Such a word is printed in italics here.

Tableau 7.8 Evaluation of output candidates of post-lexical phonology: A word-initial /d/ realized as [t] after an obstruent.

heb daar /hɛP *dar*/ 'have there'	DEPIO(VOICE)	MF(VOICE) IS(VOICE)	NVOC	IF (voice), IHS(voice)
hɛPdar			*!	
☞ hɛPtar				*
hɛPTar		*!		
hɛptar	*!			*

In conclusion, we propose an analysis based on the assumption that the voiceless realization of some initial /d/s after obstruents is an instance of phonologized hypo-articulation. It is assumed that the relevant /d/s are phonologically voiceless after obstruents so that they satisfy the phonological constraint NVOC. The relevant /d/s can satisfy NVOC, whereas other initial stops cannot, because they belong to words that can be hypo-articulated to some extent. The members of the constraint family IDENT(VOICE) on the segments of these words are ranked low in the constraint hierarchy.

Section 7.5.2.1 argued that initial fricatives are also phonologically voiceless after obstruents so as to satisfy NVOC. The overall analysis therefore assumes that the voiceless realization of underlyingly voiced fricatives as well as the voiceless realization of some word-initial /d/s after obstruents result from the high ranking of the same constraint, i.e. from the same process. It therefore generalizes over onset fricatives and initial /d/s. Previous analyses do not make this generalization, or express it with the counter-intuitive assumption that some word-initial /d/s are actually fricatives (§3.4.5).

The analysis proposed here incorporates the following two hypotheses:

- *Independence of [voice]-specification of preceding coda obstruent*
  The realization of word-initial /d/s is independent of the phonological [voice]-specification of the preceding coda obstruent. This implies that the analysis is compatible with the CNH.

- *Type of /d/-initial words which can be realized with [t]*
  Words which can be realized with devoiced initial /d/s should comprise those which are highly frequent and generally contribute little to the

propositional content of the utterance, as a speaker can typically reduce these words without running the risk of being misunderstood by the listener.

In combination with the CNH, the analysis incorporates the following hypothesis.

- *The realization of a coda obstruent before a word-initial [d]*
  The realization of a word-final obstruent before a /d/-initial word is determined by the characteristics of the /d/-initial word. If the /d/-initial word is prone to hypo-articulation, its /d/ tends to be realized as voiceless, and the preceding obstruent is consequently voiceless as well. Otherwise, the initial /d/ is voiced, and the preceding obstruent is also voiced in the majority of cases.

7.5.4 An analysis for the regular past-tense morpheme

The alveolar stop of the regular past-tense morpheme is the final obstruent of which the realization has been claimed to depend on the specifications of the preceding coda obstruent. This stop is voiceless after underlyingly voiceless obstruents, and voiced after all other types of phonemes, including underlyingly voiced obstruents (§3.4.6). Its realization must be determined in phonology, as it is related to the underlying [voice]-specification of the preceding obstruent.

In order to account for the realization of the stop after obstruents, we adopt Booij's (1995) claim that the stop is underlyingly unspecified for [voice]. We assume that it is linked to the underlying [voice]-specification of the preceding segment, if this segment is an obstruent (cf. Grijzenhout 1999). Preceding obstruents are unspecified for [voice] in the output of lexical phonology, since they are in coda position. Hence, their underlying [voice]-specifications are only accessible at the lexical level, and the stop of the past-tense morpheme must be linked to the underlying [voice]-specification of the preceding obstruent as early as at the lexical level.

In section 7.4.5, we adopted the constraint ranking for the lexical level which is repeated in (17) for convenience.

(17) Constraint ranking at the lexical level:
*CODA(VOICE) >> MF(VOICE)

In the absence of additional assumptions, this ranking ensures that the unspecified stop of the past-tense morpheme is linked to the underlying [voice]-specification of the preceding coda obstruent. It is linked so that it avoids a violation of MAXFEATURE(VOICE). This is illustrated in Tableau 7.9, in which "T" indicates an alveolar stop unspecified for [voice].

Tableau 7.9 The evaluation of output candidates of lexical-phonology: The initial stop of the past-tense morpheme after an underlyingly voiceless obstruent.

waste /ʋɑs+Tə/ 'washed' [-voice]	*CODA(VOICE)	MF(VOICE)
$(\text{ʋɑs})_\sigma(\text{Tə})_\sigma$ [-voice]	*!	
$(\text{ʋɑS})_\sigma(\text{Tə})_\sigma$		*!
☞ $(\text{ʋɑS})_\sigma(\text{tə})_\sigma$ [-voice]		

Constraint hierarchy (17) does not have the effect of linking the stop to the [voice]-specification of the preceding segment if this segment is a non-obstruent, i.e. a sonorant. The [voice]-specifications of non-obstruents, if assumed to be present in the input, are also present in the output. The initial stop of the past-tense morpheme therefore does not need to be linked to their [voice]-specification in order to avoid a violation of MAXFEATURE(VOICE).

It may be assumed that if the initial stop of the past-tense morpheme does not follow an obstruent, it receives the default specification for [voice], as it must be specified for [voice] at the lexical level. The default [voice]-specification is context-specific, and [+voice] for obstruents following sonorants and preceding vowels. This assumption can be formalized with a constraint which requires segments to be specified for [voice], in combination with a constraint which bans voiceless obstruents between sonorants. *CODA(VOICE) must dominate the former constraint so that coda obstruents can be unspecified in the output of the lexical level of phonology. IDENTSTOP(VOICE) must dominate the latter constraint, so that underlyingly voiceless obstruents in intervocalic positions can be voiceless in the output of the lexical level.

In conclusion, we assume that the initial stop of the regular past-tense morpheme is unspecified for [voice]. It is linked to the underlying [voice]-specification of the preceding obstruent, if present, as a result of the constraint ranking which expresses the CNH. After non-obstruents, the stop receives the default [voice]-specification of the relevant context.

7.5.5 Conclusions

Sections 7.5.2 to 7.5.4 presented a new analysis for the realizations of fricatives, word-initial /d/s, and the initial stop of the regular past-tense morpheme. This analysis is based on a functional explanation of why initial fricatives behave differently from most initial stops, and why only the initial /d/ of function words can be realized as [t]. It explains why initial /d/s and initial underlyingly voiced fricatives are realized as voiceless in the same context. This analysis is not based on unfounded or counter-intuitive assumptions about, for instance, the presence of word-boundaries, and is compatible with the CNH.

7.6 Conclusions

This chapter presented a new analysis of the realization of obstruents as voiced or voiceless in Dutch. It was argued that a new analysis was necessary, since previous analyses, which assume that the realization of all obstruents is completely determined in phonology, cannot explain all data.

The analysis presented in this chapter assigns an important role to phonetics. It adopts the Complete Neutralization Hypothesis, which states that coda and word-final obstruents are unspecified for [voice] in the phonological and phonetic component, and that they are realized as voiced or voiceless depending on which realization requires no additional articulatory effort, i.e. best meets the speaker's natural tendency to expend as little articulatory effort as possible. The Complete Neutralization Hypothesis can explain all available data on the realizations of neutral obstruents in Dutch.

Under the Complete Neutralization Hypothesis, the realization of onset obstruents after coda obstruents cannot result from [voice]-assimilation. The analysis therefore incorporates a new description for the realization of the initial obstruents that are realized as voiced in some contexts and as voiceless in others. It assumes that the realizations of these obstruents are completely determined in phonology, and it accurately accounts for the voiceless realizations of obstruents in the same context with the same process. It is based on a function explanation for the question why fricatives differ from stops in their realizations, and why only the segments of function words can be realized unfaithfully to their underlyingly specifications in some cases.

The overall analysis incorporates several hypotheses. Among them are the following:

Hypothesis I
Obstruents which are lexically in coda positions are realized as voiced or voiceless independently of their underlying [voice]-specifications (Hypothesis 1 in §7.4.6). In addition, they are realized as voiced or voiceless independently of the phonological feature specifications of the adjacent segments, provided that the realization of these specifications does not happen to influence acoustic characteristics which are related to the perception of voicing (see Hypotheses 2 in §7.4.6.)

Hypothesis II
Obstruents which are lexically in coda position are more likely to be realized as voiced than onset obstruents which are phonologically voiceless, and less likely to be realized as voiced than onset obstruents which are phonologically voiced (Hypothesis 4 in §7.4.6).

Hypothesis III
There is a systematic difference in realization between word-final obstruents before vowel-initial clitics in word-combinations which are likely to be retrieved as single units from the lexicon and in word-combinations which are usually computed from their parts (Hypothesis 6 in §7.4.6).

Hypothesis IV
The realization of a word-final obstruent before a /d/-initial word is determined by the tendency of the /d/-initial word to be hypo-articulated, i.e. by its frequency of occurrence, and its contribution to the propositional content of the utterance. If the /d/-initial word is prone to hypo-articulation, its /d/ tends to be realized as voiceless after obstruents, and the preceding obstruent is consequently voiceless as well. Otherwise, the initial /d/ is voiced, and the preceding obstruent is also voiced in the majority of cases (§7.5.3).

The following chapters will evaluate these hypotheses on the basis of our corpus. Chapter 8 will describe the research method, while chapters 9 and 10 will describe the actual evaluation of the hypotheses on the basis of single intervocalic stops and intervocalic geminates.

8 General research method

8.1 Introduction

Chapter 7 presented a new model of [voice] in Dutch, and a list of hypotheses implied by this model. The following two chapters will test *Hypotheses I* to *IV*, which are repeated below for convenience.

> *Hypothesis I*
> Obstruents which are lexically in coda positions are realized as voiced or voiceless independently of their underlying [voice]-specifications, and the phonological feature specifications of the adjacent segments, provided that the realization of the latter specifications does not interfere with the perception of voicing.
>
> *Hypothesis II*
> Obstruents which are lexically in coda position are more likely to be realized as voiced than onset obstruents which are phonologically voiceless, and less likely to be realized as voiced than onset obstruents which are phonologically voiced.
>
> Hypothesis *III*
> There is a systematic difference in realization between word-final obstruents before vowel-initial clitics in word-combinations which are likely to be retrieved as single units from the lexicon and in word-combinations which are usually computed from their parts.
>
> *Hypothesis IV*
> The realization of a word-final obstruent before a /d/-initial word is determined by the tendency of the /d/-initial word to be hypo-articulated, i.e. by its frequency of occurrence, and its contribution to the propositional content of the utterance.

This chapter will discuss the obstruents which will form the basis for the testing (§8.2), and the classification of these obstruents as voiced or voiceless (§8.3). In addition, it will describe the relevant statistical tests (§8.4).

8.2 The data

8.2.1 Introduction

The hypotheses to be tested concern all types of obstruents in several types of contexts. The testing will be restricted to intervocalic alveolar and bilabial stops in some types of words and word-combinations, and to intervocalic alveolar geminates. The choice of these obstruents will be motivated in section 8.2.2. The testing is further restricted in that it will not be based on all tokens of these obstruents. The tokens which are left out of the data set will be discussed in section 8.2.3. The numbers of rejected tokens and the numbers of remaining tokens, which form the initial data set, will be listed, and discussed briefly, in section 8.2.4.

8.2.2 Type of data

8.2.2.1 General type of data

Hypotheses I to IV concern the obstruents listed in Table 8.1.

Table 8.1. Obstruents relevant to each hypothesis.

Hypothesis	Obstruents
I	A. Underlyingly voiced and voiceless neutral obstruents.
	B. Neutral obstruents adjacent to segments with various specifications.
II	Obstruents which are lexically in coda position and obstruents which are lexically in onset position.
III	Word-final obstruents before vowel-initial clitics in word-combinations that may be retrieved from the lexicon, and in combinations that are not retrieved from the lexicon.
IV	Word-final obstruents before all types of /d/-initial words.

These hypotheses will not be tested on the basis of all relevant obstruents, as this is impossible both for reasons of time and limitations of the corpus. They will be evaluated only on the basis of the obstruents listed in Table 8.2. Table 8.2 is therefore the instantiation of Table 8.1 for the present study.

Table 8.2. Obstruents forming the basis for the testing of the hypotheses .

Hypothesis		Obstruents
I	A	• Post-vocalic word-final obstruents before unaccented vowel-initial function words (e.g. *had ik* /hɑd̠ ɪk/ 'had I', *zet ik* /zɛt̠ ɪk/ 'put I'). • Post-vocalic word-final /d/ and /t/ before words starting with a /d/ and an unaccented vowel (e.g. *goed denk* /xud dɛŋk/ 'good think', *dat doen* /dɑt̠ dun/ 'that do').
	B	Post-vocalic word-final obstruents before unaccented vowel-initial function words (see *I*A).
II		Post-vocalic word-final obstruents before unaccented vowel-initial function words (see *I*A). Post-vocalic word-medial obstruents before unstressed vowels (e.g *meter* /met̠ər/ 'metər').
III		Post-vocalic word-final obstruents before unaccented vowel-initial function words (see *I*A).
IV		Post-vocalic word-final /t/ and /d/ before all types of words starting with a /d/ and an unaccented vowel (see *I*B).

Hypothesis II will be evaluated on the basis of obstruents in intervocalic positions, since it is principally in these positions that neutral and non-neutral obstruents are expected to be realized differently. Neutral obstruents in truly intervocalic positions are predicted to be realized as voiced in fast speech and as voiceless in slow speech (see Hypothesis 5 in §7.4.6). Hence, in normal conversations they are predicted to be sometimes voiced and sometimes voiceless. In contrast, non-neutral obstruents are expected to be realized invariably in accordance with their [voice]-specifications.

The testing will be restricted to word-final obstruents followed by unaccented vowel-initial function words and word-medial obstruents followed by unstressed vowels so that the neutral and non-neutral obstruents that will be compared differ minimally as to syllabic and prosodic position. Unaccented function words often behave as enclitics in Dutch, and incorporate into the preceding prosodic word. They form unstressed syllables with the preceding neutral obstruents, which are consequently in the onset position of an unstressed syllable at the surface, like (non-neutral) word-medial obstruents preceding unstressed vowels.

Since *Hypothesis II* is tested on the basis of obstruents before vowel-initial clitics which are post-vocalic, the obstruents before vowel-initial clitics which will be considered for the testing of *Hypothesis III* are post-vocalic as well. Basing several tests on the same obstruents entails that no more data have to be collected than is strictly necessary. Moreover, if several tests are based on the same data set, the results of operations needed for one test may provide information that simplifies the other tests. The test procedure of *Hypothesis III*, for instance, will show which word-combinations are most likely to be retrieved as units from the lexicon. These combinations may then be omitted when testing *Hypothesis II.* As a consequence, the word-final obstruents which will be considered for the evaluation of *Hypothesis II* will all be lexically in coda position, as they should.

Hypothesis IV will be tested on the basis of obstruent clusters consisting of a post-vocalic alveolar stop and a word-initial /d/ (i.e. geminates). This type of cluster was chosen because it is represented by a sufficiently large number of tokens in the corpus. Moreover, the two alveolar stops are usually realized as one (long) stop (§3.5), which means that the cluster can be classified as voiced or voiceless by the same method as single intervocalic stops, which form the basis for the testing of the other hypotheses. This study will be restricted to geminates followed by unaccented vowels, because this type of alveolar geminates is best represented in the corpus.

Finally, *Hypothesis I* will be tested on the basis of all word-final obstruents which form the basis for the testing of the other hypotheses: intervocalic word-final obstruents and post-vocalic alveolar stops before pre-vocalic, word-initial /d/. These obstruents form an adequate data set for the testing of *Hypothesis I* and their selection allows several hypotheses to be tested on the same data set, which is economical, and simplifies the testing procedure (see above).

We will not consider all intervocalic single obstruents. First, the testing will be restricted to bilabial (/b/, /p/) and alveolar (/t/, /d/) stops. Fricatives and the velar stop will be disregarded because the voiced fricatives are phonemes for only few speakers recorded for the corpus (§5.4.3), and the voiced velar stop (/g/) for none of them (§3.3). The fricatives and the velar stop therefore cannot falsify *Hypothesis I* which states that underlying [voice]-specifications do not influence the realization of neutral obstruents. Moreover, *Hypothesis II* does not hold for these obstruents, since they cannot be misinterpreted as voiced phonemes, and are therefore probably more or less equally likely to be realized as voiced if a voiced realization requires less articulatory effort than a voiceless realization when they are lexically in coda position as in onset position.

Second, the testing of *Hypotheses I* to *III* will be restricted to stops in certain types of words and word-combinations. This is the subject of sections 8.2.2.2 and 8.2.2.3.

8.2.2.2 Types of word-combinations with intervocalic stops

The testing of *Hypotheses I* to *III* will not be based on post-vocalic word-final stops preceding all types of unaccented vowel-initial function words. Obstruents preceding rare function words will not be considered, since these obstruents can only contribute to the evaluation of a hypothesis if they are taken together as one group (i.e. pooled) with obstruents preceding other function words. Pooling is not always possible since obstruents preceding different function words often differ in many respects, such as in their position in the prosodic structure of the utterance and in segmental context. These differences can influence their realizations.

The word-combinations that will be considered are the following highly frequent ones.

(1) The word-combinations to be tested
- verb form + *ik* /ɪk/ 'I'.
  Examples: *weet ik* /ʋe$\underline{\text{t}}$ ɪk/ 'know I', *heb ik* /hɛ$\underline{\text{b}}$ ɪk/ 'have I'.
- *dat ik* /dɑ$\underline{\text{t}}$ ɪk/ 'that I'.
- verb form + *het* /ət/ 'it', with *het* having direct object function.
  Examples: *had het* /hɑ$\underline{\text{d}}$ ət/ 'had it', *hoop het* /ho$\underline{\text{p}}$ ət/ 'hope (1sg.) it'.
- verb form + *er* /ər/ 'there'.
  Examples: *zet er* /zɛ$\underline{\text{t(-t)}}$ ər/ 'put(-s) there', *liep er* /li$\underline{\text{p}}$ ər/ 'walked (sg.) there'.
- *dat er* /dɑ$\underline{\text{t}}$ ər/ 'that there'.
- *met een* /mɛ$\underline{\text{t}}$ ən/ 'with a'.

Several remarks have to be made with regard to this list. First, verb form + *ik*, *dat ik*, verb form + *het*, verb form + *er*, *dat er*, and *met een* will be referred to as combination categories. Each combination category verb form + *ik*, verb form + *het*, and verb form + *er* represents combination types in which the variable "verb form" is filled in. Examples of combination types are the examples mentioned in (1). Realizations of combination types are tokens. For instance, the realization [wetɪk] uttered in the fourth minute of the conversation between Subjects A and B is a token of /ʋet ɪk/, which is a type of the combination category verb form + *ik*.

Second, list (1) shows that the evaluation will be based on the categories *dat ik*, and *dat er*, in addition to the categories verb form + *ik*, and verb form + *er*. The category *dat het* will not be considered in addition to the category verb form + *het*, because it is represented by only a very small number of tokens in the corpus.

Third, the alveolar stop preceding *het* and *er* may represent a stem-final alveolar stop plus the verbal suffix /t/ (second or third person singular present tense). The parts of this alveolar geminate are not separated by a word-boundary, and the geminate therefore corresponds to a single segment in the output of phonology (§3.5). Under the Complete Neutralization Hypothesis (§7.4.2.3), this geminate is unspecified for [voice] in the output of phonology, since both its parts are in coda position at the lexical level and therefore unspecified for [voice]. Hence, the underlying alveolar stop geminates have the same duration and [voice]-specification as the word-final single alveolar stops in the output of phonology, and they will therefore be pooled in this study.

Fourth, the function word *er* has a lexical variant starting with /d/ (/dər/). This means that not all alveolar stops preceding [ər] in combinations of *dat er* and verb form + *er* are single and completely unspecified for [voice]. Some of them are geminates of which the second part is phonologically specified for [voice]. This part is sometimes specified as [+voice] and sometimes as [−voice], since the /d/ of /dər/ can be realized as voiceless after obstruents, like the initial /d/ of most function words (§7.5.3 and §10.2.1). The partly specified geminates are preferably not pooled with the single, completely unspecified stops for the testing of the hypotheses. Nevertheless, the two types of stops will be pooled in the present study because geminates tend to be realized as single stops, which means that the two types of stops are almost impossible to distinguish on the basis of the phonetic form. Since they are pooled, the realizations of verb form + *er* with alveolar stops, and *dat er* have to be interpreted with care. These combination categories are incorporated into this study in spite of this disadvantage because their stops are in segmental contexts similar to those of word-medial stops and therefore allow comparison with them (§9.4.3), and because few quantitative studies have been devoted to them.

Finally, *er* is translated with 'there', although not all tokens of *er* that will be examined here can actually be translated as such. See also Model (1991: 299), who argues that there are at least five types of *er* in Dutch. This study will not make a distinction between the various types, however, because none of them have such high frequencies in our corpus that they can be considered separately, and slight differences in function are not expected to influence realization (cf. §10.2.2).

8.2.2.3 Types of words with intervocalic stops

The testing of *Hypothesis II* will not be based on all tokens of word-medial, post-vocalic obstruents before unstressed vowels in the corpus, since the corpus contains several thousands of them. The investigation will be restricted to tokens of post-vocalic stops before tautomorphemic schwas, as in the words *open* /o<u>p</u>ən/ 'open' and *ladder* /lɑ<u>d</u>ər/ 'ladder', and to tokens of post-vocalic stops before the infinitive marker -/ən/, as in the verb forms *hebben* /hɛ<u>b</u>-ən/ 'have-inf.' and *weten* /ʋe<u>t</u>-ən/ 'know-inf.'. These tokens may form a data set which is sufficiently large for the evaluation of *Hypothesis II*.

Moreover, in order to avoid investigating more tokens of stops than strictly necessary, the investigations will be restricted to maximally nine tokens of bilabial and nine tokens of alveolar stops realized by each subject in each of the two contexts. These numbers were realized by nearly all subjects. We selected stop tokens in word types which maximally differ in phonological make-up, in frequency of occurrence, and in meaning. Moreover, we selected maximally two tokens of a word type realized by the same subject, and word tokens from different periods in the recording sessions. The resulting data set can be considered to be representative of the population of intervocalic stops followed by tautomorphemic schwas and the infinitive /-ən/ in the corpus.

8.2.3 Disregarded tokens

Not all tokens in our corpus can be used as data for the present study. The following categories of tokens will be left out of the data set, since they do not form reliable data for the testing of hypotheses on the realization of segments in casual speech.

1. Tokens realized simultaneously with background noise. The transcription of such tokens is difficult, if not impossible.
2. Tokens repeated several times in succession. Repeated stretches of speech are less likely to be produced in the same way as single realizations (§4.3.4).
3. Tokens in quoted speech. When quoting, a speaker sometimes imitates someone else's grammatical system. Quoted speech is therefore not representative of the speaker's own system, and quoted speech in the corpus could be unrepresentative for the general language system of the

group of speakers of the corpus.

4. Tokens that are realized while the speaker is laughing, as laughter influences speaking.

The following tokens will be disregarded as well. They may form reliable data for the evaluation of other hypotheses, but are inadequate for the testing of *Hypotheses I* to *IV*.

5. Tokens in which the intervocalic stops are separated from the following vowels by glottal stops or hesitations, or in which the vowel is absent. The stops in these tokens are not realized as truly intervocalic.
6. Tokens of word-combinations in which both the word-initial /d/ and the preceding stop were realized with release bursts. The stop sequences in these tokens were realized as obstruent clusters, instead of single stops. Since obstruent clusters and single (long) stops have different cues to the voiced/voiceless distinction, the classifications of obstruent clusters and single stops as either voiced or voiceless are not comparable, and should not be pooled for the testing of a model which assigns an important part to phonetics.

8.2.4 Numbers of rejected and accepted tokens

Table 8.3 lists the number of tokens in the corpus for each word combination category that will be considered in the investigations, the number of tokens that cannot be incorporated into the initial data set for reasons mentioned in section 8.2.3, and the number of tokens which were left, and form the initial data set. The category of intervocalic word-medial stops is not incorporated into the table, since we did not investigate all its tokens in the corpus (§8.2.2.3).

The table shows that all combination categories are represented by at least 35 tokens in the initial data set. The category with the word-initial /d/ and the category verb form + *ik* are particularly well represented, and probably form a solid basis for the evaluation of the hypotheses for that reason.

On average, 34% of the tokens could not be incorporated into the data set for reasons discussed in section 8.2.3. If the corpus had not been tape-recorded in a soundproof room, this percentage would have been even higher. Apparently, the common use to tape speech for phonological and phonetic research in a soundproof room is worthwhile.

Table 8.3. Numbers of relevant stops in the corpus, the absolute and relative (%) numbers of stops that had to be excluded form the initial data set, and the absolute and relative (%) numbers of stops that were included. The numbers are broken down by combination category.

Combination category	Numbers of stops					
	in the corpus		excluded		included	
word + /d/-initial word	1234	(100%)	363	(29%)	871	(71%)
verb form + *ik*	588	(100%)	234	(40%)	354	(60%)
dat ik	157	(100%)	71	(45%)	86	(55%)
verb form + *het*	130	(100%)	52	(40%)	78	(60%)
verb form + *er*	102	(100%)	30	(29%)	72	(71%)
dat er	56	(100%)	21	(38%)	35	(62%)
met een	86	(100%)	34	(40%)	52	(60%)

Most tokens that had to be discarded because of background noise were realized when the subjects shuffled their papers, played with their plastic cups, touched their microphones, or spoke simultaneously. Perhaps this implies that the subjects should have been more urgently requested to make as few noises as possible and to avoid interrupting each other. Such a request, however, would have decreased their spontaneity, and therefore the naturalness of their speech, and other, partial, solutions are possible: instructions written on cardboard, paper cups, etc.

The tokens in the initial data set were included in the final data set if they could be classified as either voiced or voiceless. This is the subject of section 8.3.

8.3 The classification of the stops

8.3.1 Introduction

Since the hypotheses concern the realizations of obstruents as voiced or voiceless, they cannot be tested before the obstruents in the data set are classified as such. Section 8.3.2 will describe the method of classification employed in this study, while section 8.3.3 will present an overview of its results. The consistency of the method is the subject of 8.3.4. Its validity is discussed in section 8.3.5, which compares its results to those of other methods.

8.3.2 The chosen method

There are roughly two methods for the classification of obstruents as voiced or voiceless: the obstruents can be classified on the basis of measurements of their acoustic properties and those of their surrounding vowels, and they can be classified by ear. The former method has two clear advantages: it is clear what the classifications are based on, and the classifications are listener-independent. Nevertheless, the obstruents will not be classified on the basis of acoustic measurements in this study, as this is impossible on the basis of our present knowledge of the acoustic cues to the voiced/voiceless distinction of intervocalic stops. It is unknown which values of the relevant acoustic properties relate to voiced stops, as opposed to voiceless stops, which is the relative significance of the different acoustic cues to the perception of voicing, and how the cues interact, i.e. which are the cue tradings, under which conditions (§7.3.4). For this reason, the classifications in the present study will be made by ear.

Auditory classifications have two serious problems: they are listener-dependent, and as mentioned in section 4.3.2, they are not always valid. These two problems are partly solved if auditory classifications are accepted only if they are arrived at by several independent trained phoneticians, as this minimizes the risk of classifications being listener-dependent, and increases the probability that they are valid. This is especially true if the judges are trained phoneticians, since their classifications are less likely to be influenced by their expectations, and, above all, if the segments are presented in such a way that the judges are unable to guess which are their lexical representations.

The classifications in this study are all based on the auditory judgements of the author, and on those of two other trained phoneticians who are native speakers of Dutch. In total, four phoneticians were involved in the classifications: the author and Phonetician A classified all stops in the initial data set, Phonetician B judged the word-medial /t/s and the stops followed by *ik*, *het*, and *er* that were realized by Subjects A to J, and Phonetician C classified all stops that were not judged by Phonetician B. We classified the stops independently from each other. Only stops that were unanimously classified as either voiced or voiceless were incorporated into the final data set. All others, i.e. those that were classified differently by the three of us and those that could not be classified at all, were left out.

The classifications were based on sound fragments containing the relevant stops and the adjacent vowels. The vowels were included in the fragments presented to the judges because they contain cues to the voiced/voicelessness distinction of the obstruents (§7.3.4). The judges did not hear the stops in larger contexts since they would then have known the lexical representations of the stops, and this might have influenced their voiced/voiceless classifications (§4.3.2, and see Ganong 1980 for data supporting the hypothesis that a judge's knowledge of lexical representations influences his auditory voiced/voiceless classifications).

Since the stops were played only with their adjacent vowels, the judges had little time to accustom themselves to the speech rate at which the stops were realized, and were unable to determine the positions of the stops in the prosodic structure of the utterances. There is a potential problem here as this could be argued to effect the validity of the judges' classifications. Speech rate and prosodic position influence the durations of stops and vowels (see e.g. Nooteboom & Slis 1969; Beckman & Edwards 1990; Cambier-Langeveld 1997), and listeners normally compensate, completely or partially, for these influences (e.g. Nooteboom 1979; Nooteboom & Doodeman 1980; Miller 1981). For instance, listeners base their classification of a vowel as tense or lax on its acoustic duration, in combination with its position in the prosodic structure of the utterance, and the rate of speech (Nooteboom 1979 and Nooteboom & Doodeman 1980). If listeners do not know the prosodic position of a segment and cannot determine the rate of speech, they cannot compensate for the effects of these factors on duration. Since duration is a cue to the voiced/voicelessness distinction (§7.3.4), this implies that these listeners classify certain stops differently from listeners who can compensate for these factors.

However, we do not expect that the short length of the sound fragments seriously affected the validity of the classifications. The judges were probably well able to compensate for the influences of all types of factors on duration, including prosodic structure, since they were presented with fragments which included the vowels adjacent to the relevant stop, and listeners interpret the duration of a segment on the basis of the duration of the adjacent segments, as is suggested by findings by Miller (1981), among others.

The fragments were played back by means of the speech analysis package *Praat* (Boersma 1996). They were fed into the computer with a sampling frequency of 48 kHz, and were cut from the surrounding speech at zero-crossings, so that they did not start or end with disturbing clicks. The judges heard all fragments through closed-ear headphones at approximately the same, comfortable volume. If a stop could not be immediately classified as voiced or voiceless, the relevant fragment was repeated several times. The repetitions were separated by pauses of several seconds, since continuous repetitions of a fragment may influence perception (e.g. Warren 1976).

To sum up, the classifications were based on the auditory judgements of three phoneticians. If all three phoneticians judged a stop as voiced, it was considered to be voiced, and if all three phoneticians classified it as voiceless, it was considered to be voiceless. In all other cases, the stop was not incorporated into the final data set. The judges heard the stops in short fragments, which were probably long enough to allow the judges to compensate for the influences of speech rate and prosodic position on duration, and short enough to prevent the judges from guessing the underlying or phonological representations of the stops form lexical preconceptions.

8.3.3 Overview of the results

Table 8.4 shows the numbers of stop tokens in the initial data set, the number of stop tokens for which we did not reach unanimous classification, and the number of tokens for which we did.

Table 8.4. Number of stops we attempted to classify and the absolute and relative (%) numbers for which we did and did not reach agreement. The numbers of the word-final stops are broken down by combination category.

Category	Numbers of stops in the initial data set that were					
	examined		not unanimously classified		unanimously classified	
word-medial	588	(100%)	71	(12%)	517	(88%)
word + /d/-initial word	871	(100%)	114	(14%)	757	(86%)
verb form + *ik*	354	(100%)	62	(18%)	292	(82%)
dat ik	86	(100%)	23	(27%)	63	(73%)
verb form + *het*	78	(100%)	23	(29%)	55	(71%)
verb form + *er*	72	(100%)	16	(22%)	54	(78%)
dat er	35	(100%)	3	(9%)	32	(91%)
met een	52	(100%)	10	(19%)	42	(81%)

The stop tokens which were not unanimously classified as either voiced or voiceless form 15% of the total, and maximally 29% of the tokens of a combination category. They therefore can probably be left out of the final data set without seriously depleting this set (cf. §4.3.2).

There were a few cases of segments being unanimously characterized as to [voice], but not unanimously classified as to manner, or place of articulation. We decided to regard all realizations of underlyingly bilabial stops as bilabial stops, and all realizations of underlyingly alveolar stops as alveolar stops.

8.3.4 The consistency of the chosen method

In order to estimate the consistency (i.e. reliability) of the judgements, we asked the judges to classify 100 stops anew which they had unanimously classified as either voiced or voiceless before. These 100 stops were representative of the stops in the different combination categories and the different speakers in the corpus. Nine of Phonetician A's classifications did not agree with his original classifications, three of Phonetician B, and six of Phonetician C. The phoneticians therefore arrived at different classifications in on average 6% of cases, which means that they were reasonably consistent. We could not test the consistency of our own judgements, since the testing took place after we had thoroughly processed the data, and our new classifications would have been very likely to be affected by the patterns which we had by that stage discovered in the data.

Since the phoneticians were reasonably consistent in their judgements, the probability is negligible that a certain stop which was unanimously classified as voiced will be classified as voiceless, or vice versa, a second time around. This probability is the multiplication of the probabilities (p_d) that every phonetician judging a certain stop arrives at a deviant classification (see for instance Ross 1976). The most plausible assumption with respect to p_d is that it is different for each judge, and approximately 0.09 for Phonetician A, 0.03 for Phonetician B, and 0.06 for Phonetician C. Under this assumption, it is impossible to compute the actual probability that a stop which was unanimously classified as voiced by ourselves and two other judges will be classified as voiceless, or vice versa, the second time around, since the probability with which we classify a stop differently ourselves is unknown. We therefore assume that each judge arrives at deviant classifications in on average 6% of cases, and that $p_d = 0.06$ for all judges. Under this assumption, the probability with which a stop will be unanimously classified differently the second time around is 0.06 * 0.06 * 0.06, which is smaller than 0.1 %, and therefore approximately zero.

The probability that a stop which is incorporated into data set should be disregarded on the basis of reclassifications is much greater. It is the probability that only one or two judges arrive at a different classification when we classify the stop the second time around. Under the assumption that $p_d = 0.06$ for every judge, this probability equals $3 * p_d * (1 - p_d) * (1 - p_d) + 3 * p_d * p_d * (1 - p_d) = (3 * 0.06 * 0.94 * 0.94) + (3 * 0.06 * 0.06 * 0.94) = 0.16 + 0.01 = 0.17$, i.e. 17%. This suggests that when a stop which was classified unanimously as either voiced or voiceless is judged a second time, it will be judged differently by one judge in on average 16% of cases, and by two judges in on average 1% of cases.

In conclusion, the method of classification employed in this study is consistent in that stops which were classified as voiced will hardly ever be reclassified as voiceless, or vice versa. The method is less consistent in that the probability is 17% that a stop which was classified as either voiced or voiceless will turn out to be unclassifiable the second time around.

8.3.5 Differences from other methods

8.3.5.1 Introduction

The validity of a classification method can be investigated by comparing its results to those of other classification methods. If there are great differences, the method is perhaps not particularly valid. We compared the classifications obtained in this study to those which may result from a method based on acoustic measurements, and to those which result when the stops are presented to the judges in fragments containing several syllables.

8.3.5.2 Methods based on acoustic measurements

It was mentioned in section 8.3.2 that the classifications to be used in this study would not be based on acoustic measurements because it is not exactly known what the different acoustic properties contribute to the voiced/voiceless opposition in the case of intervocalic obstruents in casual speech. It might be possible to develop a classification method based on acoustic measurements given the characteristics of the stops that were unanimously classified as voiced or voiceless in the present study. If this is possible, the classifications obtained will relate to a number of acoustic characteristics, and can certainly be considered to be valid.

Section 7.3.4 discussed all cues to the voiced/voicelessness distinction that have been discovered so far. Some of them are not highly relevant to intervocalic stops, e.g. the characteristics of the formant transitions. Others can only be measured with difficulty, such as the intensity of the noise burst, and are therefore not sufficiently reliable as a basis for classifications. Cues which are probably highly relevant and can be more or less accurately measured are the durations of the closure and the noise burst of the stops, they might perhaps serve as the basis for a classification method.

In order to investigate this possibility, we investigated the durations of the closures and bursts of the single alveolar stops in word-medial and word-final intervocalic positions in the final data set. These durations were measured by means of the speech analysis software *Praat* (Boersma 1996). First, the approximate positions of the vowels, and the closure and the burst of the stop were determined in the waveform. This was done by ear, and by visual inspection of spectrograms, intensity trace curves, and the form of the signal in the waveforms (i.e. amplitude signals). Next, the positions of the boundaries were determined in the waveforms on the basis of the form of the signal. A decrease in the height of the regular waveform was assumed to indicate that the closure is made, while the start of irregular waveforms with very short periods indicates that the closure is released, and the burst has begun. A decrease in the height of these irregular waveforms indicates the end of the burst. All boundaries found on the basis of the waveforms were checked by ear. As an example, Figure 8.1 (page 204) shows the waveform and the spectrogram of part of the [ɛi], the [t], and part of the [ə] of *in feite* /ɪn fɛitə/ 'actual', realized by Subject L. Vertical lines indicate the boundaries between the vowel and the closure, between the closure and the burst, and the end of the burst. Finally, the intervals between the boundaries were measured.

We measured the closure and burst durations of 649 stops, and found that several stops have bursts which are longer than 54 ms. Since such long bursts are believed to be uncommon in careful speech, this suggests that bursts may be longer in casual than in careful speech. In addition, it was found that several stops were not realized with complete closures. We assumed that these stops have closures which correspond to the interval in which the intensity of the sound is lower. These intervals are not marked as clearly in the signal as complete closures, and the durations assigned to these stops are therefore not as valid as the durations assigned to stops with complete closures. Figure 8.2 (page 205) shows the waveform and intensity trace of a stop realized with an incomplete closure, and parts of the adjacent vowels. The boundaries of the closure of the stop are indicated again with vertical lines. The stop comes from a token of *had ik* /hɑd ɪk/ 'had I''realized by Subject M as [hɑd ɪk].

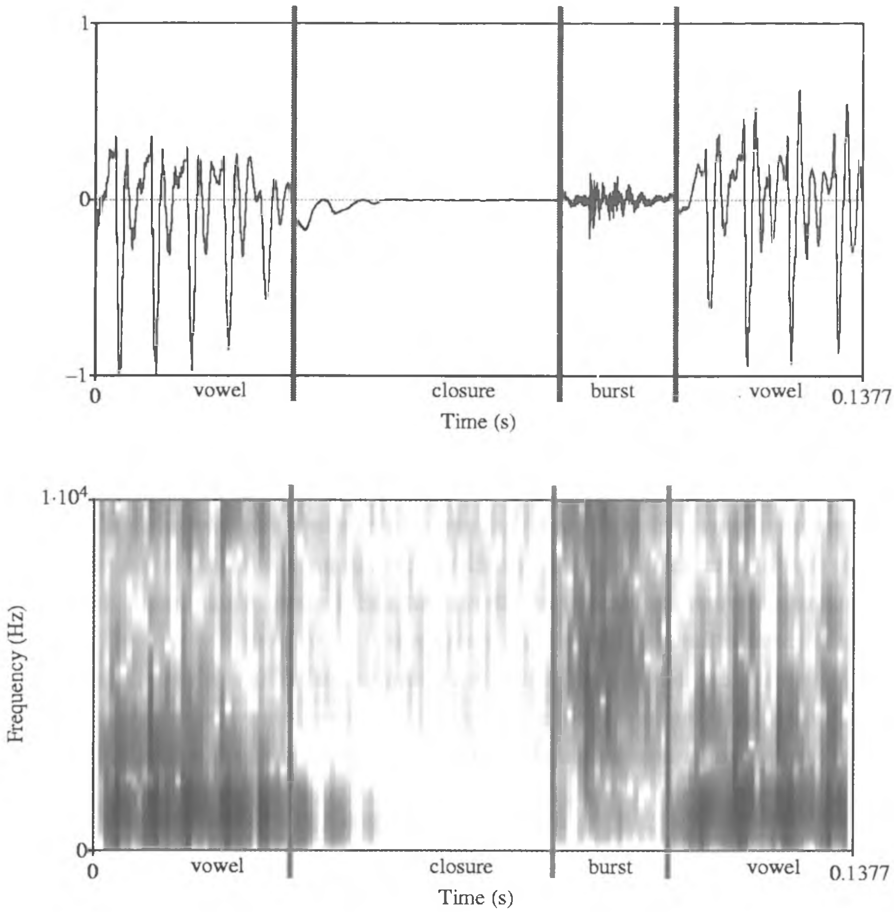


Figure 8.1 Closure and burst boundaries in the waveform and spectrogram (analysis width = 1ms; time step = 1ms) of [ɛitə].

In order to estimate the consistency of the measurements, the durations of 100 randomly chosen closures and 100 randomly chosen bursts were determined twice. The average difference between two measurements of one and the same duration appeared to be 2.8 ms in the case of the closures, and 2.7 ms in the case of the bursts. In 62% of cases, the measurements of the same duration differed 2 ms or less. When interpreting the durations, these differences should be taken into account.

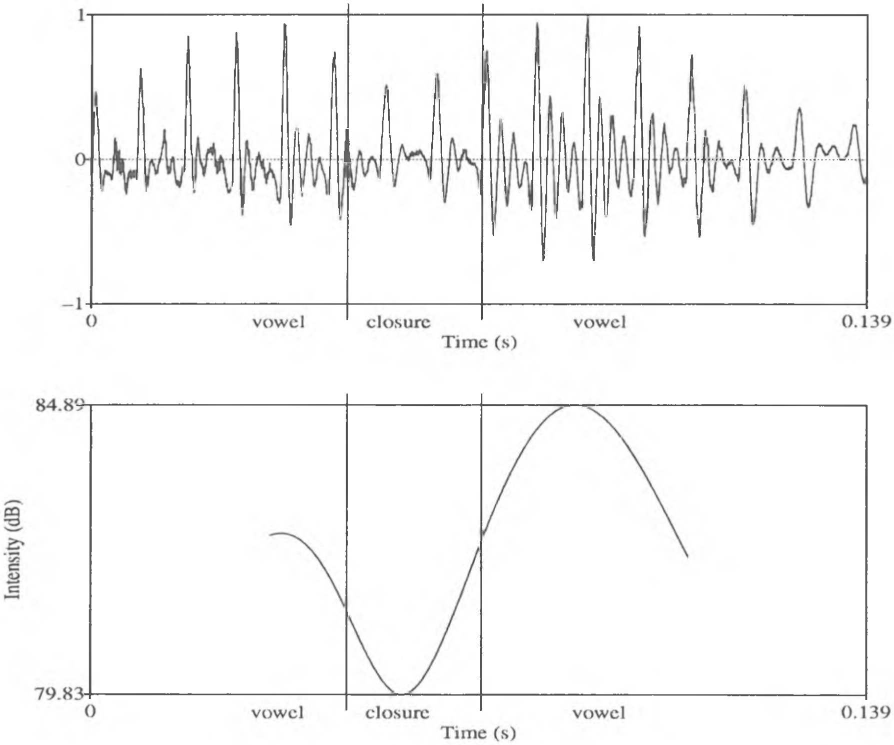


Figure 8.2 Closure boundaries in the waveform and intensity trace (time step = 1 ms) of [ɑdɪk].

Figures 8.3 shows the closure durations of the stops that were classified as voiced, while Figure 8.4 shows the closure durations of the stops that were classified as voiceless. Figures 8.5 and 8.6 show the burst durations of these stops. All four histograms indicate that stops with short durations tend to be classified as voiced, whereas stops with long durations are generally perceived as voiceless. This suggests that closure and burst durations are valuable predictors of the realization of intervocalic stops as voiced or voiceless in casual speech.

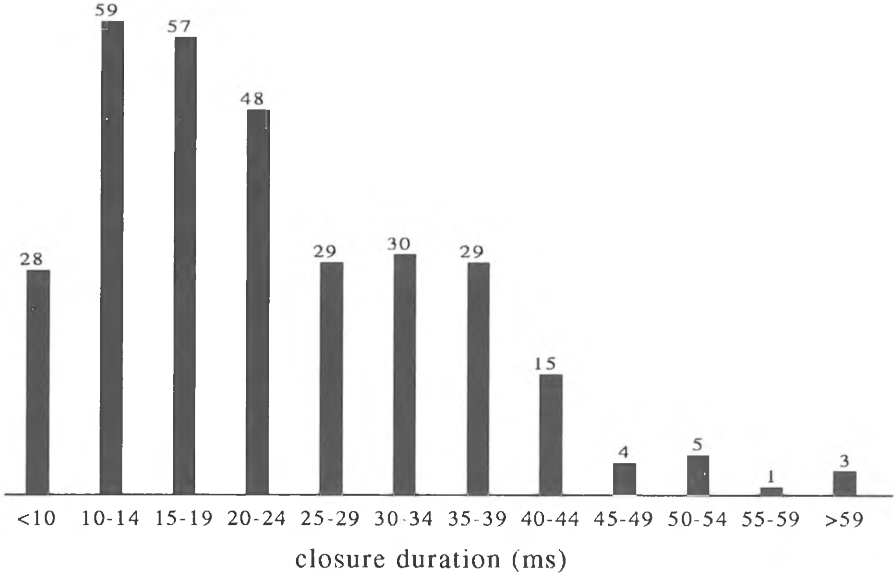


Figure 8.3 The closure durations of the voiced alveolar stops.

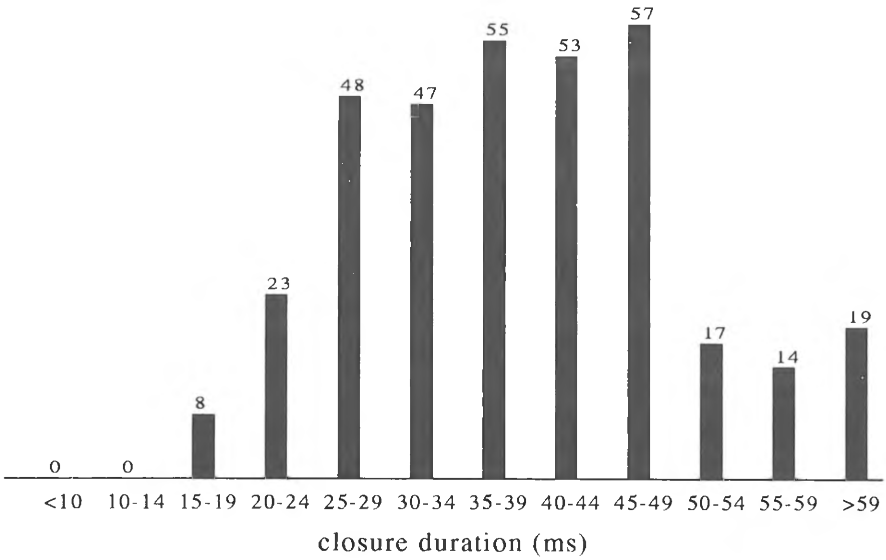


Figure 8.4 The closure durations of the voiceless alveolar stops

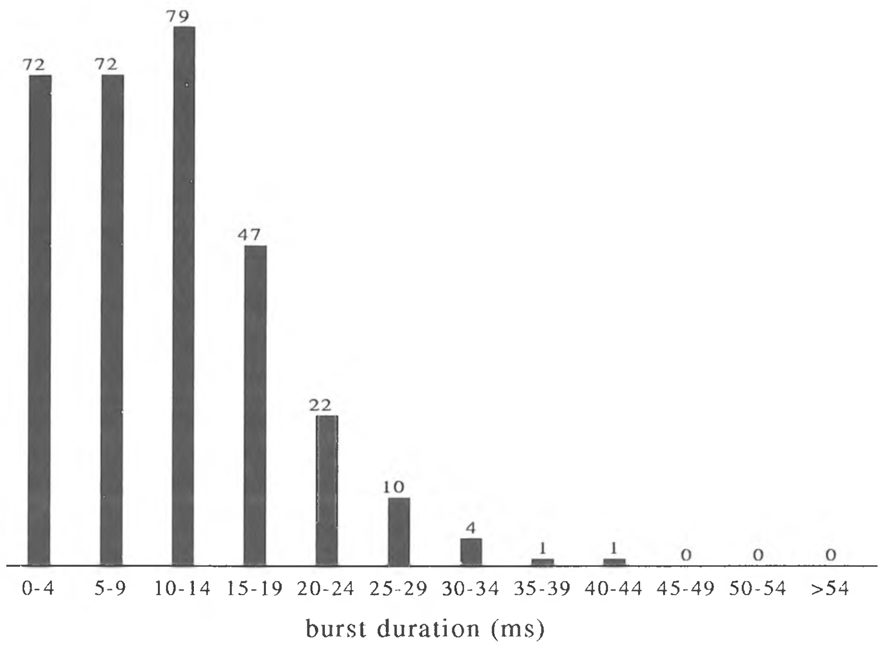


Figure 8.5 The burst durations of the voiced alveolar stops.

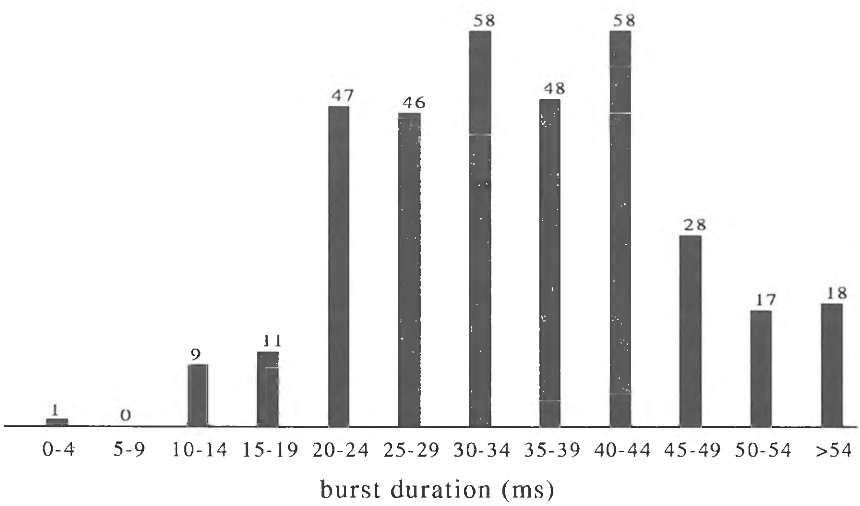


Figure 8.6 The burst durations of the voiceless alveolar stops.

Figure 8.7 is a plot of the closure durations of the stops against their burst durations. Voiced stops are indicated as d-s, and voiceless ones as t-s. The indicated line separates the t-s from the d-s as well as possible. It was calculated with the method described in Appendix F, and accounts for 93.5% of the classifications. Its slope indicates that the classification of a stop is somewhat more strongly related to its burst duration than to its closure duration.

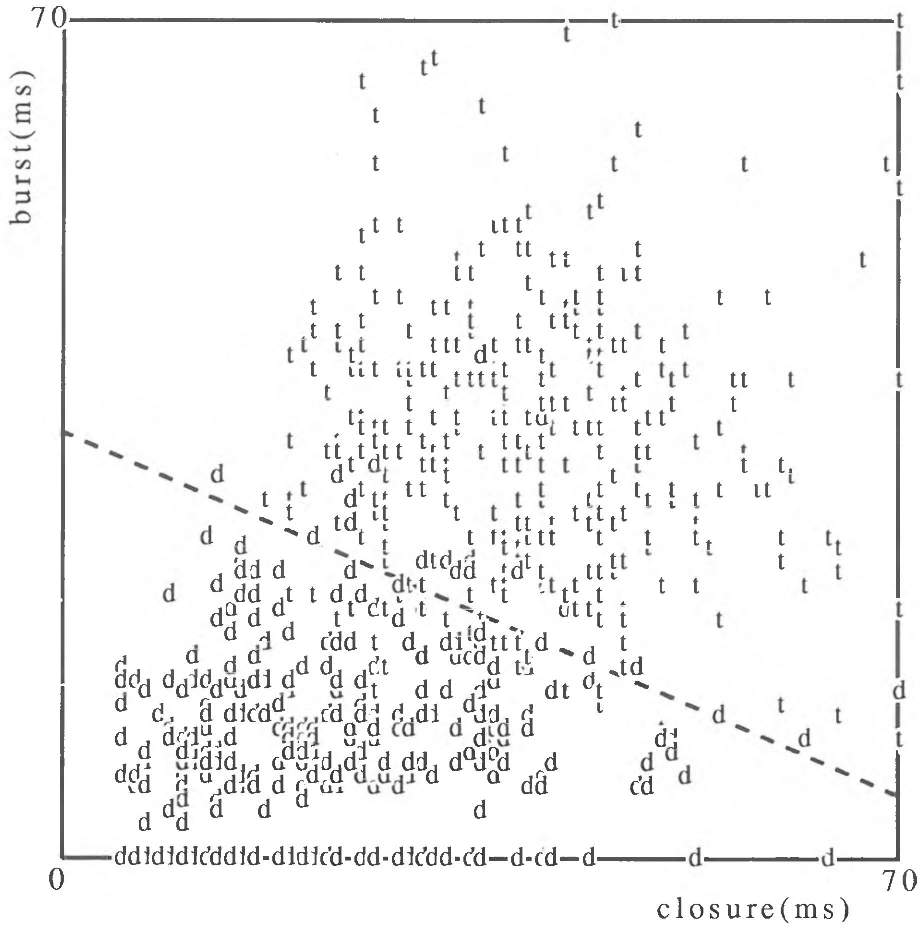


Figure 8.7 The closure and burst durations of the voiced (d) and voiceless (t) alveolar stops.

Since there is an overlap of 93.5% between the results of the auditory classifications and the results which would have been obtained if the classifications were based on the line in Figure 8.5, the acoustic measurements largely confirm the judges' classifications of the stops as voiced or voiceless. They support the validity of the auditory classifications.

Given these results, it could be argued that the classification method based on the auditory judgements of human listeners can now be replaced by a method based on the closure and burst durations of the stops, but this is not quite true, since the classifications made by ear are based on more cues than closure and burst duration. As will be shown in section 9.2, these additional cues cause the exact relation between the classification of a stop and its durations to be different for different contexts. The relation cannot be calculated for every context on the basis of the present data, since they are too few in number. This means that the stops can at present only be classified on the basis of the relation which holds for the whole data set, which certainly produces inaccurate classifications for stops in certain types of contexts. Therefore, classifications based on the closure and burst durations of the stops cannot be highly valid as yet, and classifications made by ear are preferable.

8.3.5.3 Auditory methods based on long fragments

The auditory classifications to be used in this study were based on fragments consisting of the relevant stop and the adjacent vowels. In which follows, we will refer to these fragments as "short fragments". In order to judge the validity of these classifications, we compared them also to auditory classifications based on fragments consisting of four syllables before, and four syllables after, if possible, the relevant stop. The latter fragments will be referred to as "long fragments".

The most obvious way to compare the results of classifications based on short and long fragments is by classifying a certain number of stops in both types of fragments, and comparing the results. If some stops are unanimously classified as voiced on the basis of the short fragments and as voiceless on the basis of the long fragments, or vice versa, the two classification methods produce different results. These differences cannot be interpreted as resulting solely from the fallibility of the classification method based on the short fragments, since we have established in section 8.3.4 that this method is consistent.

Comparing the two methods in this way was impossible in the present study, since it meant that not only the three phoneticians but also we ourselves would have to classify a number of stops a second time. This would have run into the same problems that we explained in section 8.3.4 above: our own reclassifications would have been all too likely to be affected by our familiarity with the data, and therefore to be incomparable to our original classifications.

The two methods were compared as follows. We asked the three phoneticians to reclassify 100 stops on the basis of the same short fragments (§8.3.4), and 100 stops on the basis of the long fragments. These 200 stops had been unanimously classified as either voiced or voiceless in the original classifications. Then we compared the number and types of stops presented in the short and in the long fragments for which each phonetician arrived at deviant classifications.

Table 8.5 shows the numbers of stops that were judged differently by each phonetician. The figures in this table show that the phoneticians' judgements were inconsistent in about the same number of stops regardless of the length of the fragments.

Table 8.5 Numbers of stops which were reclassified differently by the phoneticians, broken down by the length of the fragments presented.

Length of the fragments	Phonetician		
	A	B	C
short	9	3	6
long	9	3	8

Independently of the lengths of the fragments, Phoneticians A and C classified differently both stops that had been classified as voiced and stops that had been classified as voiceless in the original classifications. Phonetician B, on the other hand, classified differently both types of stops only when they were represented in the short fragments: all stops that she classified differently in the long fragments were originally classified as voiced. The probability that three stops which were originally classified as voiced are judged as voiceless if stops which were classified as either voiced or voiceless are just as likely to be judged differently is $(0.5)^3 = 0.13$, i.e. 13%, and therefore above chance-level. This suggests that the fact that Phonetician B's deviant judgements only involved stops which were originally classified as voiced may be due to chance.

In conclusion, the data suggest that classifications based on short fragments hardly differ from those based on long fragments. This suggests that the auditory classifications to be used in the investigations can be considered valid.

The phoneticians disagreed on the ease with which stops are classified on the basis of the short and long fragments. Phoneticians A and C thought that classifying stops in short fragments is easier. They had difficulties in concentrating on the relevant stop in the long fragments, while phonetician C apparently experienced the additional problem of being influenced by his own expectations when he heard several syllables, and could guess what type of word or word-combination he was dealing with. Phonetician B, on the other hand, preferred the long fragments to the short ones. She could probably compensate more easily for the influences of the speech rate and prosodic structure on duration when hearing several syllables.

8.3.6 Summary

In conclusion, three phoneticians classified every stop by ear on the basis of fragments containing the stop at issue and the adjacent vowels. The fragments presented were sufficiently short not to reveal the lexical representations of the stops, and sufficiently long to allow the judges to benefit from all acoustic cues related to the voiced/voiceless distinction, and to compensate for the influences of the speech rate and prosodic structure on duration. A stop is assumed to be either voiced or voiceless if the phoneticians unanimously classified it as such. In total, 85% of the stops were classified unanimously. The classification method appeared to produce consistent results, since the probability that a certain stop is sometimes unanimously classified as voiced and sometimes as voiceless was found to be less than 0.1%. In addition, the method produces valid results, since the classifications are scarcely affected by the length of the fragment presented, and are closely related to the closure and burst durations of the stops.

8.4 Statistical analysis of the data

The preceding sections noted that 34% out of the relevant word and word combination tokens cannot be taken into account because they do not constitute useful, reliable data, and that 15% out of the remaining tokens cannot be incorporated into the final data set because their stops could not be classified as either voiced or voiceless. The hypotheses formulated in section 7.6, therefore, can be tested only on the basis of the remaining 56% relevant tokens.

These tokens form a sufficiently large data set only if the tokens realized by the different subjects in the corpus are pooled. If they are not pooled, the sample sizes will be too small. There are two facts which indicate that pooling is legitimate. The first one is that the subjects speak similar varieties of Standard Dutch (§5.4.3), which implies that their grammars are roughly identical. The second one is that the data set does not suggest that the speakers differ in the realization of obstruents as voiced or voiceless, which, unfortunately, cannot be confirmed by studies of inter-speaker variability as the data set is, again, too small. For these reasons, the tokens realized by the different subjects will be pooled in the present study.

The hypotheses will be tested using Kendall's rank-correlation test (see e.g. Liebetrau 1983), Fisher's exact test (see e.g. Siegel & Castellan 1988: 103), and log-linear analyses (see e.g. Rietveld & van Hout 1993).

Kendall's rank-correlation test indicates whether there is a relation between two variables which are (approximately) continuously distributed, such as the frequency of occurrence of a word, and the proportion of tokens realized with voiced stops. Fisher's exact test indicates whether two groups of tokens, for instance word-medial and word-final stops, differ in the proportions with which they are assigned to two classes, such as realized as voiced or voiceless. Both Kendall's test and Fisher's exact test do not suppose certain distributions for the dependent variables, and are therefore non-parametric. A further advantage is that they can be applied to small numbers, which is important for this study.

A log-linear analysis can detect the relations between frequencies of occurrence (dependent variable) and one or more categorical, independent variables. It also detects the interactions among the independent variables. It can, for instance, detect the relations among the relative number of times a stop is perceived as voiced (dependent variable), the underlying [voice]-specification of the stop (independent variable), and the roundness of the preceding vowel (independent variable). Since log-linear analyses cannot handle cells containing zeros, 0.5 is added by default to each cell. This addition does not significantly affect the relative differences among the cells, and therefore the results of the log-linear analyses, if the cells which do not contain zeros contain relatively high values. In cases in which both a log-linear analysis and Fisher's exact test can be applied, we will report the results of only one test. The other test was run as well, and found to provide the same type of results.

The hypotheses that will be tested are of two types. *Hypotheses II* to *IV* are of the conventional type, stating that the independent variable has an effect. We will use a significance level of 5% for the tests of these hypotheses. *Hypothesis I* is of a different type, since it states that there will be no effect of the independent variable. A significance level of 5% is too low for this type of hypothesis, since, if an effect is not found to be significant at the 5% level, and the associated p-level is 8%, it is still quite likely that the relevant variable has an effect in the population. We will assume that the data do not disconfirm *Hypothesis I* only if $p > 0.20$.

If it cannot be determined with a certainty of at least 80% that a certain phonological variable has an effect, this will of course only significantly affect

Hypothesis I if the data set is very large. If the data set is small, such a finding cannot be taken as evidence that there is no effect, because it may be the case that the data set is simply too small to reveal it. The significance of a finding that a certain sample shows no effect is expressed by the power of the test, which is the probability of finding a statistically significant effect in the sample if the populations from which the sample is drawn differ in the variable at issue (see e.g. Rietveld & van Hout 1993). We will not calculate powers in the present study, since all tests in this study are based on relatively small samples, which implies that the probability is always large that no statistically significant effect is found, even if the variable at issue actually does have an effect. An additional reason why powers will not be calculated is that *Hypothesis I* concerns the effects of phonological variables. Phonological variables have categorical effects by definition (§2.4.1), and therefore, if they affect voicing, they generally cause large differences between populations representing their different values, provided that these populations contain approximately the same type of tokens. This implies that it is not the case that each effect, i.e. each difference between populations, no matter how small, is relevant for the testing of *Hypothesis I*. The significance of the finding that $p > 0.2$ should not be expressed by the probability that *some* difference is nevertheless present between the populations, as powers do, but by the probability that a *large* difference is present.

In order to determine the significance of a finding that $p > 0.2$, we will estimate the maximal difference that may be present between the populations representing the values of the phonological variable at issue. This maximal difference can be estimated by means of formula (2), provided that the distribution of the dependent variable can be approximated by a normal distribution. This is the case in the present study if the sample sizes are larger than 25, as the dependent variable "proportion of voiced stops" has a binomial distribution. If the difference between two populations turns out to be small, we will assume that the effect of the relevant phonological variable is absent.

(2) *Estimation ($p < 0.05$, two-tailed) of the maximal difference (d_{max}) between the proportions of tokens representing one value of the dependent variable (e.g. classified as voiced) in two populations which represent the values of the independent variable (e.g. underlyingly voiced and voiceless). The symbols* n_1 *and* n_2 *represent the size of the two samples, and* p_1 *and* p_2 *the proportion in these samples. The samples should be numbered such that* p_1 *is larger than, or equal to,* p_2.

$$d_{\max} = p_1 - p_2 + 1.96 \sqrt{\frac{p_1 (1 - p_1)}{n_1} + \frac{p_2 (1 - p_2)}{n_2}}$$

It is unfortunately unknown which is the minimal difference that may be due to a phonological factor. This probably depends on the factor itself, and on the composition of the data sets.

8.5 Conclusions

This chapter described the stops which will form the basis for the testing of the hypotheses, the classification of these stops as voiced or voiceless, and the statistical analyses that will be run on them. The most important findings are that the auditory classification method used in this study is highly consistent and probably valid, that auditory classifications of intervocalic stops realized in casual Dutch as voiced or voiceless are closely related to the closure and burst durations of the stops, and finally, that many tokens of relevant stops in the corpus cannot be used because they were realized with too much background noise. The latter finding confirms the general assumption that recordings for phonological/phonetic research should be made in a soundproof room.

9 Single intervocalic stops

9.1 Introduction

This chapter will evaluate *Hypotheses I* to *III* (§7.6) on the basis of the single stops in the data set (§8.2.2). First, section 9.2 will present the overall data, and show that they do not obviously falsify the Complete Neutralization Hypothesis, which implies that the detailed testing of the hypotheses is worthwhile. Then, *Hypotheses I* to *III* will be tested one after the other in sections 9.3 to 9.5.

The hypotheses are independent of each other, and therefore can be tested in any order. The testing procedure of *Hypothesis III*, however, will determine which word-combinations behave exceptionally, and therefore which ones should not be considered for the evaluation of *Hypotheses I* and *II*. This implies that *Hypothesis III* is preferably tested before *Hypotheses I* and *II*, and so this was done. The evaluation of *Hypotheses I* and *II* will nevertheless be presented first, because they are more crucial for the analysis.

9.2 Overall data

Appendix G lists the numbers of voiced and voiceless single stops in the data set. The word-final stops are ordered by category and type of word-combination, while the word-medial ones are ordered by word type. Table 9.1 presents a summary for the word-final stops.

Table 9.1 Absolute and relative (%) frequencies of occurrence of voiced and voiceless word-final stops, broken down by word-combination category.

Combination category	Numbers of stops classified as			
	voiced		voiceless	
verb form + *ik*	252	(86%)	40	(14%)
dat ik	27	(43%)	36	(57%)
verb form + *het*	35	(64%)	20	(36%)
verb form + *er*	8	(14%)	48	(86%)
dat er	0	(0%)	33	(100%)
met een	14	(33%)	28	(67%)

The table shows that nearly all combination categories are represented by tokens with voiced and by tokens with voiceless intervocalic stops. This observation is in accordance with the Complete Neutralization Hypothesis.

In addition, the table shows that the combination categories differ in their proportion of tokens with voiced stops. One might argue that these differences falsify the Complete Neutralization Hypothesis, since this hypothesis states that word-final stops are unspecified for [voice] in all types of combination. All word-final stops are specified identically, and therefore should be realized identically. If this reasoning is correct, the testing of *Hypotheses I* to *IV* is not worthwhile.

The reasoning is flawed, however, since the differences among the combination categories may be due exclusively to non-phonological factors. They can, for instance, at least in part be ascribed to the differences in frequency with which the types of the different categories are retrieved as units from the lexicon. The types of some combination categories, such as those containing *dat* (§9.5), are more likely to be retrieved as units from the lexicon than other categories, and are therefore more likely to have word-final stops which are in onset position at the lexical level of phonology. As a consequence, their word-final stops are more likely to be specified for [voice] in the input of phonetics, and to be systematically realized as either voiced or voiceless.

The differences may further be due to, for instance, a difference in the proportion of tokens with bilabial stops. Section 8.3.5.2 showed that the phoneticians' classifications are related to the burst and closure durations of the stops. This does not exclude the possibility that their judgments were (also) based on other cues to voicing. The presence of glottal vibration before the release of the stop may have played a role, when the closure of the stop was long or had a medium length (cf. §7.3.4). The cavity above the glottis is larger for bilabial than for alveolar stops, and the larger surface area of its walls can more easily expand under high air pressure (Ohala 1983). Consequently, more air can pass the glottis in bilabial stops than in alveolar stops before the vocal folds stop vibrating. Hence, if the presence of glottal vibration plays a role, long bilabial stops are more likely to be perceived as voiced than long alveolar stops.

Finally, we would like to suggest that the differences between the combination categories may well be due to differences in the height of the vowels preceding the stops. High vowels are realized with smaller oral cavities and larger pharyngeal volumes than low vowels (see for instance the X-ray figures in Fant 1960, and the MRI-pictures in Rietveld & van Heuven 1997: 77). While the shape of the oral cavity changes during the realization of an alveolar stop, the volume of the pharynx remains more or less constant. Alveolar stops following high vowels are therefore realized with the same oral volume as alveolar stops following low vowels, but with a larger pharyngeal volume. As a consequence, more air can be accommodated in the vocal tract during the realization of alveolar stops preceded by high than by low vowels, and the vocal folds can vibrate longer (Ohala 1983). Hence, if the presence of glottal vibration during the last part of the closure affects the classifications of long stops, long stops sound as voiced with longer closures when preceded by high vowels than when preceded by low ones.

We evaluated this possibility that the height of the preceding vowel determines the closure duration with which stops are perceived as voiced. Figure 9.1 plots the burst and closure durations of all alveolar stops following [+high] vowels, while Figure 9.2 plots those of all alveolar stops following [−high] vowels. The voiced stops are represented as *d*-s, the voiceless ones as *t*-s. Each plot contains a separation line (printed as a fat dashed line), which optimally separates the voiced from the voiceless stops. It presents the relation between the classifications of the stops and their durations. The thin dashed line in Figure 9.1 represents the separation line of Figure 9.2, while the thin dashed line in Figure 9.2 represents the separation line of Figure 9.1. The lines in a figure show that the separation line for stops following [+high] vowels is positioned above the separation line for stops following [−high] vowels, and that this is particularly the case for medium and long closure durations. The difference between the separation lines is statistically significant (see Table 9.2 for the characteristics of the lines, and Appendix F for the method of calculation and comparison of the lines). The data therefore suggest that stops with the same burst durations are classified as voiced with longer closure durations when they are preceded by high vowels than when they are preceded by low ones, and that the difference is especially important for stops with medium and long closures. The height of the preceding vowel apparently affects the voiced/voiceless classification, which implies that differences between the combination types may be due to differences in the percentages of their tokens with long stops following high vowels.[1]

Table 9.2 Characteristics of the lines separating voiced and voiceless alveolar stops preceded by high or low vowels, on the basis of their burst and closure durations. Definitions of the characteristics can be found in Appendix F.

Separation line for stops following	Slope	Standard error slope	Position	Standard error position
high vowels	65.6°	2.4°	35.3 ms	1.5 ms
low vowels	59.9°	3.6°	32.5 ms	1.4 ms

[1] Remark that Figures 9.1 and 9.2 cannot be explained by the assumption that, when listeners classify stops as voiced or voiceless on the basis of duration, they compensate for the fact that the bursts of stops following high vowels tend to be shorter than the bursts of stops following low vowels, which is a result of the difference in the size of the vocal tract. If listeners compensated for this difference, they would consider stops following high vowels as voiceless with shorter burst durations than stops following low vowels. This turns out to be contrary to fact.

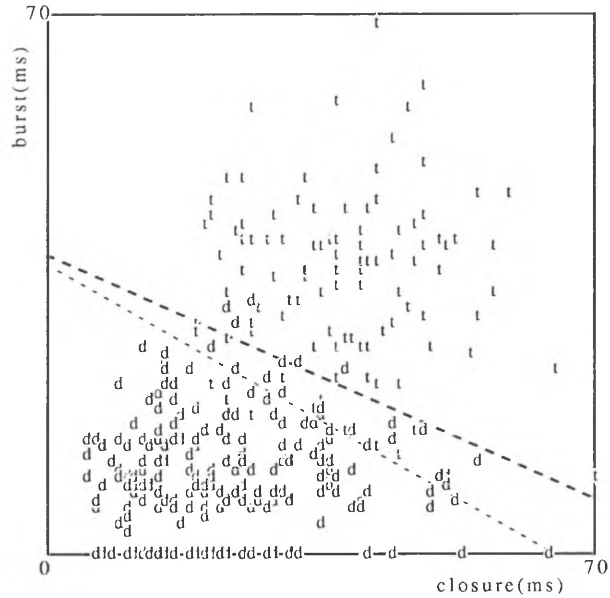


Figure 9.1 The closure and burst durations of the voiced (d) and voiceless (t) alveolar stops following [+high] vowels.

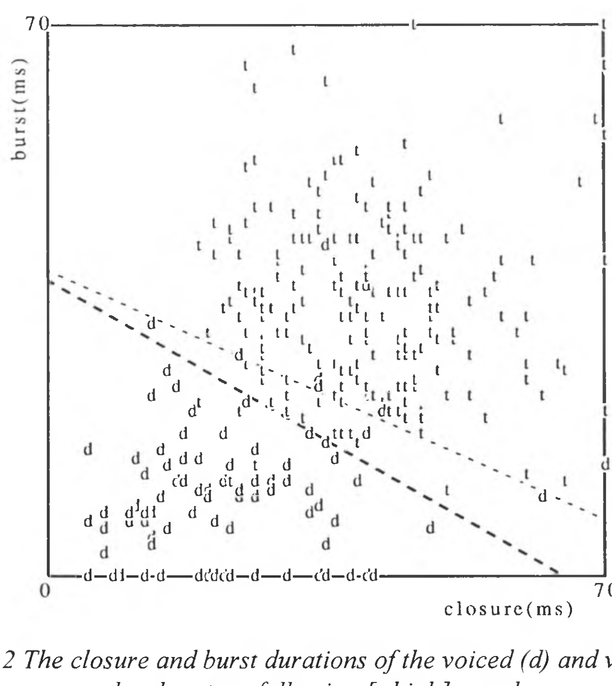


Figure 9.2 The closure and burst durations of the voiced (d) and voiceless (t) alveolar stops following [-high] vowels.

In conclusion, differences among the types present no argument against the Complete Neutralization Hypothesis. The testing of *Hypotheses I* to *IV* is therefore worthwhile.

9.3 The effect of feature-specifications on neutral stops

9.3.1 Introduction

9.3.1.1 Variables to be tested

The first hypothesis that will be tested is *Hypothesis I*. It is repeated below for convenience.

> *Hypothesis I*
> Obstruents which are lexically in coda positions are realized as voiced or voiceless independently of their underlying [voice]-specifications, and the phonological feature specifications of the adjacent segments, provided that the realization of these latter specifications does not interfere with the perception of voicing.

The evaluation of this hypothesis can, and will, consist of a test of the influence of the phonological features listed in (1).

(1) The features of which the influences will be studied for the evaluation of *Hypothesis I*
- the underlying feature [voice] of the neutral stop (§9.3.2);
- the feature [round] of the preceding vowel (§9.3.3);
- the feature [lax] of the preceding vowel (§9.3.4).

Therefore, the phonological features of the following vowel as well as the features of the stop and the preceding vowel that are not mentioned in (1) will be left out of consideration. These features are listed in (2).

(2) The features of which the influences will not be studied for the evaluation of *Hypothesis I*.
- the place of articulation of the stop;
- the feature [high] of the preceding vowel;
- the feature [back] of the preceding vowel;
- the features of the following vowel.

As was argued in section 9.2, the realization of the articulation place of a stop as well as the realization of the height of the preceding vowel may influence acoustic characteristics which are related to the perception of voicing. As a consequence, any difference in classification between bilabial and alveolar stops, and between stops following high and low vowels may be purely phonetic in nature, and is meaningless with respect to *Hypothesis I*. The influence of these features will therefore not be investigated for the testing of this hypothesis.

The feature [back] is related to the feature [high]: most front vowels are high, and most back vowels are low in Dutch (see Table 3.1). Since a difference in classification between stops following high and low vowels may be purely phonetic in nature, a difference between stops following front and back vowels may be purely phonetic as well. Therefore, the influence of this feature will also be left out of consideration.

Finally, this study will not deal with the type of vowel following the stop. This vowel is always /ɪ/ or /ə/ (§8.2.2.2), which means that its phonological properties vary only minimally in the tokens. Another reason is that it is difficult to determine which stops are not followed by /ə/, as the only function word with /ɪ/, i.e. *ik* /ɪk/ 'I', is believed to have a lexical variant with /ə/ (/ək/) (Berendsen 1986: 36; Booij 1995: 167), and the /ɪ/ and the unspecified vowel /ə/ can be realized alike between an alveolar or bilabial stop and a velar stop.

9.3.1.2 Data

The evaluation of *Hypothesis I* will be based on the word-final stops in the combination types of verb form + *ik*, verb form + *het*, and verb form + *er*, except *heb ik* /hɛb ɪk/ 'have I'. The reason why the combination *heb ik* and tokens containing *dat* are left out of consideration is that they are very likely to be retrieved as single units from the lexicon (§9.5), which implies that the final stops of their first words are very likely to be lexically in onset, instead of coda, position The combination *met een* is not considered because all its tokens contain an underlyingly voiceless stop as well as a non-round lax vowel, and the influence of the features in (1) therefore cannot be tested on them.

For the evaluation of the hypothesis, the different combination categories are not pooled, since the stops in these combinations are followed by different function words, and therefore by different segments. This difference may well influence the phonetic implementation of the stops, and therefore the exact frequency with which they are realized as voiced and voiceless.

9.3.1.3 Method

For the evaluation of the influence of the features listed in (1), we will compare voiced stops to underlyingly voiceless ones, stops following round vowels to stops following non-round vowels, and stops following lax vowels to those following tense vowels.

Because long stops tend to be more often classified as voiced when they are bilabial and follow high vowels (§9.2), the groups of stops which are compared should preferably comprise similar percentages of long bilabial stops as well as similar percentages of long stops following high vowels. We did not test whether this is the case for all groups of stops that are compared, because the classifications of the word-final stops in the relevant combination categories do not seem to be so strongly related to the articulation places of the stops and the heights of the preceding vowels that they cannot show the effects of phonological factors. Table 9.3 shows that the bilabial stops were not more likely to be classified as voiced than the alveolar ones in these categories (Fisher's exact test $p > 0.05$ one-tailed for all combination categories). These categories probably contain too few stops which are sufficiently long to show the influence of articulation place. The type *heb ik* is not incorporated into this table, since it behaves exceptionally (§9.5.2). Table 9.4 shows the numbers of voiced and voiceless stops following high and low vowels in the relevant combination categories. The stops following diphthongs are not included, since diphthongs are partly [-high] and partly [+high] (§3.2). The figures in this table do show a difference between the stops following high and low vowels, but this difference is statistically significant only for the tokens of verb form + *het* (Fisher's exact test < 0.05 one-tailed), and the difference between the two classes of tokens of this combination category seems to be mainly due to the behaviour of *had het* /hɑd hət/ 'had it'. We hypothesize that the height of the preceding vowels has some effect on the classifications of the stops as voiced or voiceless, but that this effect is generally sufficiently small not to obscure the effects of phonological factors on the classifications. In conclusion, we will pool all alveolar and bilabial stops and all stops following high and low vowels in the relevant combination categories, as this may be assumed to be legitimate in an investigation of the influence of phonological variables.

Table 9.3 Absolute and relative (%) frequencies of occurrence of voiced and voiceless word-final stops, broken down by word-combination category, and place of articulation.

Combination category	Place of articulation	Numbers of stops classified as voiced		voiceless	
verb form + *ik*	alveolar	124	(80%)	31	(20%)
	bilabial	11	(69%)	5	(31%)
verb form + *het*	alveolar	28	(63%)	16	(37%)
	bilabial	7	(63%)	4	(37%)
verb form + *er*	alveolar	6	(14%)	36	(86%)
	bilabial	2	(14%)	12	(86%)

Table 9.4 Absolute and relative (%) frequencies of occurrence of voiced and voiceless word-final stops, broken down by word-combination category, and by the height of the preceding vowel.

Combination category	Vowel height	Numbers of stops classified as voiced		voiceless	
verb form + *ik*	[+high]	95	(81%)	23	(19%)
	[-high]	34	(74%)	12	(26%)
verb form + *het*	[+high]	25	(76%)	8	(24%)
	[-high]	8	(40%)	12	(60%)
verb form + *er*	[+high]	4	(18%)	19	(82%)
	[-high]	4	(13%)	28	(87%)

If the comparison of the classifications of two groups of stops does not reveal an effect of a certain phonological feature, this does not imply that this feature does not affect the acoustic properties which are cues to the voiced/voiceless distinction at all, and that *Hypothesis I* is correct. The classifications are mainly related to the burst and closure durations of the stops (§8.3.5.2), and the feature at issue may particularly affect the other cues to voicing. Its influence on the classifications may therefore only surface under special circumstances, e.g. when the stops are realized very carefully, or in certain contexts. For the testing of *Hypothesis I*, one should therefore not only consider the relations between the features and the classifications, but also the relation between the features and all acoustic properties, including or excluding burst and closure durations, which can be related to the perception of voicing.

If the data are taken from a corpus of spontaneous speech, such an investigation cannot consist of a direct comparison of all acoustic characteristics of the stops representing the different values of the feature. The characteristics of segments are influenced by many factors, such as loudness, speech rate, and position in the prosodic structure, and these factors do not have the same value for all segments realized in spontaneous speech. For instance, some segments are louder and realized at a lower speech rate than others. This means that if the data come from spontaneous conversations, the relevant acoustic characteristics can only be compared after the influence of a number of factors has been compensated for. This is a difficult task, since there is still no model available that accounts for the influences of all relevant factors. Another way out is to consider so very many tokens that the influences of the different factors are cancelled out. The required number of tokens for this, however, is probably huge, and higher than the number of stops that are provided by most existing corpora. For this reason, the influence of a feature on all acoustic properties cannot be determined on the basis of spontaneous speech by a direct comparison of all acoustic properties of the stops.

It can be determined on the basis of such speech, however, if the relation between the relevant phonological feature and the classifications of the stops is investigated, and the influence of the durations on the classifications is compensated for. Such an investigation focuses on the influence of the feature on the relation between the durations and the classifications, instead of on the classifications themselves. If stops which differ neither in their durations nor in their specification for the phonological feature turn out to be classified differently, the phonological feature affects acoustic properties related to the perception of voicing, and *Hypothesis I* is invalid.

The problem with this investigation method is that the relation between the classifications of the stops and their durations is influenced by the contexts of the stops, such as the height of the preceding vowel (§9.2). The influence of a feature on the acoustic characteristics related to the voiced/voiceless distinction may therefore be investigated on the basis of the relation between classifications and durations only if stops in different contexts are considered separately, or if the sample is so large that differences are levelled out. This implies that the data set should consist of many stops in exactly the same context or of a huge number of stops in different contexts, which is not the case in the present study.

In conclusion, the influence of the different features on acoustic properties other than burst and closure durations of the stops cannot be investigated in the present study. We will therefore test the influences of the underlying [voice]-specifications of the stops, and the roundness and the laxness of the vowels on the realization of neutral stops only by examining their influence on the voiced/voiceless classifications.

The influences of the features on the classifications will not be studied simultaneously, because this would require a log-linear analysis, and the available amount of data is insufficiently large for such an analysis. The features will be considered one by one, and the influence of each of them will be tested by a "simple" log-linear analysis with two independent variables (specification of the feature and combination category), and an estimation of the maximal difference in the proportion of voiced stops between the two populations representing the two values of the feature at issue (§8.4).

9.3.2 The effect of underlying [voice]-specifications on neutral stops

The influence of the underlying [voice]-specification of the stops on their classifications is tested on the basis of combinations ending in *ik* and *het*. The tokens containing *er* are left out of consideration, because they nearly all have underlyingly voiceless stops.

Table 9.5 shows the number of underlyingly voiced and the number of underlyingly voiceless stops in the tokens of the two combination types that were perceived as voiced and voiceless. It shows that both the underlyingly voiced and the underlyingly voiceless stops were often classified as voiced.

Table 9.5 Absolute and relative (%) frequencies of occurrence of voiced and voiceless stops, broken down by combination category and underlying [voice]-specification.

Combination category	Underlying specification	Numbers of stops classified as	
		voiced	voiceless
verb form + *ik*	[+voice]	31 (79%)	8 (21%)
	[−voice]	104 (79%)	28 (21%)
verb form + *het*	[+voice]	13 (59%)	9 (41%)
	[−voice]	22 (67%)	11 (33%)

The data do not falsify *Hypothesis I*. They would if they suggested that the probability is larger than 80% that the underlying [voice]-specifications influence the classifications (§8.4). A log-linear analysis with "the frequencies of classification as voiced or voiceless" as the dependent variable and "underlying [voice]-specification" and "combination category" as independent variables shows that neither the variable "underlying [voice]-specification" ($z = 0.435$, $p > 0.2$) nor the interaction between "underlying [voice]-specification" and "combination category" ($z = 0.465$, $p > 0.2$) is significant, if the critical $p = 0.2$. *Hypothesis I* can therefore be maintained. Only the variable "combination category" has a statistically significant effect ($z = 2.233$, $p < 0.05$), possibly for reasons mentioned in section 9.2.

Moreover, if there is a difference in the population between the proportions of underlyingly voiced and voiceless stops before *ik* that are perceived as voiced, it is small. With a probability of 0.95, it is maximally 15% (formula 2 in §8.4). Differences of this size are sufficiently small to be due to accidental differences among the tokens. They can, for instance, be due to a difference in the proportion of tokens with underlyingly voiced and voiceless stops that are retrieved as single units from the lexicon. On the other hand, they are probably too small to be due to an effect of the underlying [voice]-specifications of the stops, as phonological features typically cause large differences between the groups representing their different values (as argued in §8.4). This means that if there is a difference, this difference is not the result of the influence of the underlying [voice]-specification of a stop on its realization.

The maximal difference between the proportions of underlyingly voiced and voiceless stops before *het* that are perceived as voiced cannot be calculated on the basis of the available data. There are fewer than 25 tokens of verb form + *het* with underlyingly voiced stops in the data set, and therefore formula (2) in section 8.4 is not applicable.

In conclusion, the classifications suggest that the realization of neutral stops is not influenced by their underlying [voice]-specifications. This is in accordance with *Hypothesis I*.

9.3.3 The effect of the roundness of vowels on following neutral stops

The second feature which is considered for the testing of *Hypothesis I* is the roundness of the preceding vowel. Table 9.6 presents the numbers of stops following round and non-round vowels in the data set which were classified as voiced and voiceless.

Table 9.6 Absolute and relative (%) frequencies of occurrence of voiced and voiceless stops, broken down by combination category and the roundness of the preceding vowel.

Combination category	Specification of the vowel	Numbers of stops classified as voiced		voiceless	
verb form + *ik*	[+round]	40	(80%)	10	(20%)
	[-round]	95	(79%)	26	(21%)
verb form + *het*	[+round]	7	(64%)	4	(36%)
	[-round]	28	(64%)	16	(36%)
verb form + *er*	[+round]	1	(9%)	10	(91%)
	[-round]	7	(16%)	38	(84%)

These data show that intervocalic neutral stops following round as well as non-round vowels are sometimes classified as voiced and sometimes as voiceless. This is in accordance with *Hypothesis I*, as this hypothesis implies that roundness is of no influence.

In order to test *Hypothesis I* precisely, a log-linear analysis with "the frequencies of classification as voiced or voiceless" as dependent variable, and "roundness" and "combination category" as independent variables was run on the data in Table 9.6. The results of this analysis show that only the variable "combination category" has a statistically significant effect on the classifications (with the following z-values for the associated parameters of this effect: $z_1 = 5.4$, $p < 0.05$; $z_2 = 1.8$, $p < 0.05$). Neither "roundness" ($z = -0.2$, $p > 0.2$), nor the interaction between "roundness" and "combination category" ($z_1 = 0.2$, $p > 0.2$; $z_2 = 0.1$, $p > 0.2$) are statistically significant (see §8.4 for the choice of the critical p). Hence, the results of this analysis do not suggest that the roundness of a vowel influences the realization of the following neutral stop as voiced or voiceless, and are in accordance with *Hypothesis I*.

The maximal difference, if there is any, between the proportion of voiced stops following round vowels and the proportion of voiced stops following non-round vowels in the population of stops preceding *ik* does not suggest an effect of roundness either. With a certainty of 95%, the difference is maximally 15% (formula 2 in §8.4), and therefore probably too small to be due to an effect of the phonological variable [round] on the classifications (§8.4 and §9.3.2). The maximal differences could not be computed for the stops in verb form + *het* and verb form +

er as these combinations are represented by fewer than 25 tokens with round vowels in the data set (cf. §8.4).

In conclusion, the data from the corpus show that there is no strong relation between the roundness of a vowel and the classification of the directly following word-final stop as voiced or voiceless. This finding is in accordance with *Hypothesis I.*

9.3.4 The effect of the laxness of vowels on following neutral stops

The final feature which is considered for the testing of *Hypothesis I* is the laxness of the vowel preceding the neutral stop. We investigated its influence on the basis of the combination categories verb form + *ik* and verb form + *er*. The category verb form + *het* is left out of consideration, since it is represented in the data set by only one combination type with a lax vowel, i.e. *had het* /hɑd ət/ 'had it'.

Table 9.7 lists the number of voiced and voiceless stops which follow lax and tense vowels in the data set. The stops following diphthongs are left out of consideration, since the tense/lax distinction is considered to be irrelevant for them. The figures in the table shows that voiced stops can follow both lax and tense vowels, which is in accordance with *Hypothesis I.*

Table 9.7 Absolute and relative (%) frequencies of occurrence of voiced and voiceless stops, broken down by combination category and the laxness of the preceding vowel.

Combination category	Specification of the vowel	Numbers of stops classified as	
		voiced	voiceless
verb form + *ik*	[+lax]	41 (77%)	12 (23%)
	[−lax]	88 (79%)	23 (21%)
verb form + *er*	[+lax]	5 (18%)	23 (82%)
	[−lax]	3 (11%)	24 (89%)

Since the null-hypothesis is that there is no effect of laxness on the realization of the following stop, the difference between the proportions of voiced stops following tense and lax vowels is statistically significant if $p \leq 0.2$ (§8.4). A log-linear analysis shows that neither the variable "laxness" ($z = -0.5$, $p > 0.2$), nor the interaction between "laxness" and "combination category" ($z = 0.8$, $p > 0.2$) have such small p-values. Consequently, the variable "laxness" cannot be assumed to influence the [voice]-classifications, which is in line with *Hypothesis I.* Only the variable "combination category" has a statistically significant effect ($z = 7.1$, $p < 0.05$), possibly for reasons mentioned in section 9.2.

If there is a difference between the proportions of voiced stops following tense and lax vowels, it is small. The difference is maximally 16% for the combinations verb form + *ik*, and 25% for the combinations verb form + *er*, with an uncertainty of 5% (formula 2, §8.4). These differences are probably simply too small to be ascribed to a phonological process, and can be assumed to be due to differences among the tokens which are not related to the opposition between tense and lax vowels.

The assumption that the laxness of the preceding vowel has no effect is supported by the fact that the tokens of verb form + *ik* were more often classified with voiced stops when containing tense than when containing lax vowels, whereas the opposite is the case for tokens of verb form + *er*. Opposite effects often indicate that the variable at issue actually has no effect.

In conclusion, the data do not show a strong relation between the laxness of a vowel and the realization of the following neutral stop as voiced or voiceless. This finding is in accordance with *Hypothesis I.*

9.3.5 Conclusions

The classifications of the word-final intervocalic stops in the corpus as voiced or voiceless are in accordance with *Hypothesis I.* As is consistent with this hypothesis, they show no relations with the underlying [voice]-specifications of the stops, or the height or laxness of the preceding vowels. If they are related to these features, the relations are weak, and probably due to phonetic or lexical factors which accidentally vary with the features.

9.4 Word-medial versus word-final stops

9.4.1 Introduction

The second hypothesis which will be tested on the basis of the single stops is *Hypothesis II*, which is repeated below for convenience.

> *Hypothesis II*
> Obstruents which are lexically in coda position are more likely to be realized as voiced than onset obstruents that are phonologically voiceless, and less likely to be realized as voiced than onset obstruents that are phonologically voiced.

The testing of this hypothesis will consist of a comparison of the word-final stops, which are lexically in coda position, to the word-medial intervocalic stops, which are in onset position. The word-final stops were considered in isolation in sections 9.2 and 9.3. The word-medial intervocalic stops will be discussed in isolation in section 9.4.2. This section will show that the realization of these obstruents is in line with the analysis proposed in Chapter 7, and that obstruents preceding tautomorphemic schwas and obstruents preceding the infinitive marker *-en* can be pooled for the purpose of comparing the word-medial stops with the word-final ones. The word-final and word-medial stops will be compared in section 9.4.3.

9.4.2 Word-medial stops

Table 9.8 lists the numbers of word-medial stops that were classified as voiced and voiceless. The numbers are broken down by the underlying [voice]-specification of the stops, and the type of morpheme which contains the following schwa.

Table 9.8 Absolute and relative (%) frequencies of occurrence of voiced and voiceless word-medial stops, broken down by underlying [voice]-specification and the type of morpheme containing the following schwa.

Underlying specification	Type of morpheme	Numbers of stops classified as voiced		voiceless	
[+voice]	stem	136	(99%)	1	(1%)
	suffix	57	(100%)	0	(0%)
[−voice]	stem	9	(6%)	149	(94%)
	suffix	10	(6%)	155	(94%)

Nearly all stops were classified in accordance with their underlying [voice]-specifications. This suggests that the underlying [voice]-specifications of the stops influence their realizations as voiced or voiceless. This conclusion is supported by a log-linear analysis with "the frequencies of classification as voiced or voiceless" as dependent variable, and "underlying [voice]-specification" and "type of morpheme" as independent variables. This analysis resulted in one significant effect: "underlying [voice]-specification" ($z = 8.6$, $p < 0.05$). The variable "type of morpheme" ($z = -0.2$, $p > 0.2$), and the interaction between "underlying [voice]-specification" and "type of morpheme" ($z = -0.1$, $p > 0.2$) are not statistically significant. Since the word-medial stops are classified in accordance with their underlying [voice]-specifications, they apparently enter phonetics with [voice]-specifications that are identical to the underlying specifications, which is in accordance with the analysis developed in Chapter 7.

The finding that "type of morpheme" has no effect is in line with this analysis as well. It also indicates that for the purposes of a comparison of the word-medial stops with the word-final ones, the word-medial stops preceding tautomorphemic schwa can be pooled with those preceding *-en*.

The great majority of word-medial stops that were not classified according to their underlying [voice]-specifications are underlyingly voiceless, and were transcribed as voiced fricatives or voiced approximants. In order to realize an intervocalic obstruent as voiceless in fast speech, the speaker must take care that he realizes the obstruent as relatively long acoustically. For a voiced realization, on the other hand, he does probably not have to expend additional articulatory effort. Hence, a possible explanation for why it was particularly the phonologically voiceless stops that were realized unfaithfully with respect to [voice] is that it is more difficult to realize phonologically voiceless than phonologically voiced stops faithfully in fast speech. The assumption that the relevant stops were realized fast is supported by the observation that they were realized with incomplete closures.

9.4.3 Word-medial and word-final stops

Hypothesis II is evaluated by means of a comparison of all word-medial stops discussed in section 9.4.2 to the word-final stops in *met een* and in the combination types of verb form + *er*. The word-medial stops are not compared to the word-final stops in the combinations verb form + *ik* and verb form + *het* because these word-final stops are all followed by a schwa and an obstruent, whereas the word-medial stops are nearly all followed by at least two sonorants. This difference may influence the phonetic implementation of the intervocalic stop, and therefore the exact frequency with which it is classified as voiced and voiceless. The word-medial stops are not compared to the word-final stop in *dat er*, because the final stop of *dat* might behave exceptionally (§9.5.4). Given the conclusions reached in section 9.4.2, the word-medial stops followed by tautomorphemic schwa and those followed by the infinitive marker will be pooled.

Table 9.9 repeats the classifications of the stops at issue. It shows that both the word-final stops in tokens of verb form + *er* and in tokens of *met een* were classified as voiced more often than the phonologically voiceless word-medial stops. It further shows that these word-final stops were classified as voiced less often than the word-medial stops which are phonologically voiced. All differences are statistically significant according to Fisher's exact test ($p < 0.05$ two-tailed). Hence, *Hypothesis II* is confirmed by the data.

Table 9.9 Absolute and relative (%) frequencies of occurrence of voiced and voiceless stops broken down by class

Class of stop	Numbers of stops classified as			
	voiced		voiceless	
word-medial [+voice]	193	(99%)	1	(1%)
word-medial [−voice]	19	(6%)	304	(94%)
in verb form + *er*	8	(14%)	48	(86%)
in *met een*	14	(33%)	28	(67%)

The word-final (neutral) stops show a closer resemblance to the word-medial stops which are underlyingly voiceless than to the underlyingly voiced ones. This is the case for the intervocalic stop in *met een* at least partly because the preceding vowel [ɛ] is low and consequently does not favour the presence of glottal vibration in long stops (§9.2). It is the case for the word-final stops in verb form + *er* because some of them were not directly followed by a vowel, but by the /d/-initial variant of *er* (/dər/). Alveolar stops followed by this variant are part of a geminate of which the second part belongs to a function word. Such clusters are often phonologically specified as [-voice], and realized as voiceless (cf. §7.5.3, and Chapter 10).

9.5 Retrieval of complete word-combinations

9.5.1 Introduction

The final hypothesis which will be tested on the basis of the single stops is *Hypothesis III*, which is repeated below.

> Hypothesis *III*
> There is a systematic difference in realization between word-final obstruents before vowel-initial clitics in word-combinations which are likely to be retrieved as single units from the lexicon and in word-combinations which are usually computed from their parts.

For the evaluation of *Hypothesis III*, one must know the relative frequencies with which word-combination types are retrieved as single units from the lexicon. The word-combination types with unpredictable properties are retrieved in 100% of cases. The frequency with which the other combination types are retrieved is probably related to their frequency of occurrence, which implies that their frequencies of occurrence are relevant to the evaluation of *Hypothesis III*.

This study assumes that the frequencies of occurrence of word-combinations types are proportional to their numbers of occurrences in the pilot study of the Eindhoven corpus of spoken Dutch (Uit den Boogaart 1975), and in our corpus (see Tables G.1 to G.3 in Appendix G). The two corpora together contain approximately 244,000 tokens of words (121,569 tokens in the relevant part of the Eindhoven corpus, and approximately 122,500 tokens in our corpus). The reason why frequencies were based on two corpora instead of one is that two corpora contain more word tokens, and cover more types of conversation topics than one corpus, which means that frequencies based on two corpora are more reliable, and that differences in frequency among word-combinations are more obvious. The reason why the Eindhoven corpus has been chosen as the second corpus is that this corpus contains spontaneous Standard Dutch, is well accessible, and in its orthographic transcription the reduced and full forms of content words are distinguished.

None of the combination types in the data set are assumed to be temporarily stored in the lexicon during the recording of the corpora (cf. §2.2.1). This assumption is based on the observation that they all seem to be approximately equally frequent in all parts of our corpus and the Eindhoven corpus. It implies that none of the word-combination types is expected to be more likely to be retrieved as a single unit from the lexicon than is indicated by its frequency of occurrence.

The hypothesis will be tested on the basis of the combinations ending in *ik* (§9.5.2) the combinations of verb form + *het* (§9.5.3), and, finally, the combinations ending in *er* or *een* (§9.5.4).

9.5.2 Obstruents in combinations ending in *ik*

Word-combinations of verb form + *ik* have no known unpredictable properties. The frequency with which they are retrieved from the lexicon is therefore probably related to their frequency of occurrence.

Their frequencies of retrieval could be assumed to be proportional to their frequency of occurrence. Under this assumption, *Hypothesis III* states that there is a relation between the more or less continuously distributed variables "proportion of voiced stops" and "frequency of occurrence". This hypothesis can be evaluated by means of Kendall's rank-correlation test (§8.4). When this test is run on the basis of the data in Table G.1 (Appendix G), which show the frequencies of occurrence of all verb form + *ik* combinations, plus on the datum that *dat ik* /dɑt ɪk/ 'that I' has a frequency of 283 and a proportion of tokens with voiced stops of 0.43, it reveals no statistically significant relation ($S = 17$, number of pairs = 26, $p > 0.05$ two-tailed).

It is possible that no relation was found because the test should be run separately on the combination types with underlyingly voiced and underlyingly voiceless word-final stops. Obstruents which are voiced in the underlying form of a word are possibly also voiced in the lexical forms of word-combinations containing this word. Consequently, when these combinations are retrieved as a unit from the lexicon, and the relevant obstruent is in onset position, they are consistently realized with a voiced stop. Similarly, combinations with obstruents which are underlyingly voiceless in the underlying forms of the separate words are perhaps consistently realized with voiceless obstruents when they are retrieved as single units from the lexicon. Therefore, when word-combinations are more likely to be retrieved as single units from the lexicon, those with an underlyingly voiced word-final obstruent could be more likely to be realized with a voiced obstruent, whereas those with an underlyingly voiceless word-final obstruent could be more likely to be realized with a voiceless obstruent. This implies that combination types with underlyingly voiced and voiceless stops should be considered separately for the evaluation of *Hypothesis III*. Running Kendall's test on both classes of combination types separately does not reveal a statistically significant relation between the proportion of tokens with voiced stops and the frequency of occurrence of the combination type for either class ($S = 0$, number of pairs $= 6$, $p > 0.05$ one-tailed for the combination types with underlyingly voiced stops, and $S = 9$, number of pairs $= 20$, $p > 0.05$ one-tailed for the combination types with underlyingly voiceless stops). This implies that, given the assumption that frequency of retrieval is proportional to frequency of occurrence, no evidence for *Hypothesis III* is found.

Hypothesis III can probably also be tested by comparing the classifications of the stored combinations to those of the combinations which are not stored. The combination types which have lexical representations are probably frequently retrieved from the lexicon, since this is the only reason why storing them could be economical. The combinations which do not have lexical representations are never retrieved as units from the lexicon.

In order to find out which combinations are stored, we studied their frequency of occurrence. Table 9.10 shows the numbers of voiced and voiceless stops in combination types ending in *ik* as a function of the frequency of occurrence of the types. When the table lists just one instance of a combination type of a certain frequency, this frequency is represented by just this one type in the data set. When two types are listed, the frequency is represented by several ones. On the basis of Table 9.10 one may conclude that *heb ik* is probably present in the lexicon, as it is highly frequent, and that *stap ik* is probably absent, as it is of a low frequency. It cannot be concluded whether, for instance, *had ik* or *zit ik* is stored in the lexicon, since the exact threshold frequency for storage is unknown.

Table 9.10 Absolute and relative (%) frequencies of occurrence of voiced and voiceless stops in combination types ending in ik, *broken down by the frequency of occurrence of the combination types in the present corpus and the pilot study of the Eindhoven corpus of spoken Dutch.*

Frequency of occurrence	Instances of the combinations		Numbers of stops classified as			
			voiced		voiceless	
354	*heb ik*	'have I'	117	(97%)	4	(3%)
283	*dat ik*	'that I'	27	(43%)	36	(57%)
148	*weet ik*	'know I'	39	(81%)	9	(19%)
102	*moet ik*	'must I'	36	(82%)	8	(18%)
89	*had ik*	'had I'	27	(79%)	7	(21%)
38	*laat ik*	'let I'	3	(75%)	1	(25%)
24	*zit ik*	'sit I'	10	(83%)	2	(17%)
19	*begrijp ik*	'understand I'	5	(100%)	0	(0%)
6 - 11	*snap ik* *deed ik*	'understand I' 'did I'	3	(75%)	1	(25%)
1-5	*stap ik* *mijd ik*	'step I' 'avoid I'	12	(60%)	8	(40%)

In order to find out which of the other word-combination types have lexical representations, we studied the possibility of contraction. Some verb form + *ik* combinations have contracted forms, whereas others have not. For instance, the combination *heb ik* /hɛb ɪk/ 'have I' can be realized as [hɛk] (e.g. Booij 1995: 178), but *stap ik* /stɑp ɪk/ 'step I' cannot be realized as [stɑk]. Since not all contractions are permissible, the permissible ones must be listed separately in the lexicon. It is probably only the word-combinations of which the full forms are stored in the lexicon because of their high frequency of occurrence that have contracted variants. If this is correct, the full variants of the contracted forms are stored in the lexicon as well.

It is yet unknown which combinations can be realized in contracted forms. Hence, before lexical storage can be studied on the basis of contractions, the combinations that have contracted variants must first be identified.

Such an investigation cannot be based on our corpus, as this corpus contains few contracted forms. We decided to base it on intuition. We asked several staff members of the Faculty of Letters at the Free University Amsterdam which types of combination in the data set they think can be realized in contracted forms. The results of this small test suggest that most speakers of Dutch think that the strings *heb ik* /hɛb ɪk/ 'have I', *dat ik* /dɑt ɪk/ 'that I', *weet ik* /ʋet ɪk/ 'know I', *moet ik* /mut ɪk/ 'must I', *had ik* /hɑd ɪk/ 'had I', *laat ik* /lat ɪk/ 'let I', and *deed ik* /ded ɪk/ 'did I' can certainly be realized as [hɛk], [dɑk], [wek], [muk], [hɑk], [lak] and [dek]. Speakers hesitate about the wellformedness of [zɪk] (< *zit ik* /zɪt ɪk/ 'sit I'), [zɑk] (< *zat ik* /zɑt ɪk/ 'sat I') , [zɛk] (< *zet ik* /zɛt ɪk/ 'put I'), and [sxɑk] (< *schat ik* /sxɑt ɪk/ 'estimate I'), or express conflicting opinions.

The contracted forms which are grammatical according to most speakers correspond to highly frequent full forms, which is in accordance with the hypothesis (see above) that full forms may have contracted variants only if they are stored in the lexicon. The full forms corresponding to these contracted forms are therefore probably stored in the lexicon. The lexical status of the full forms corresponding to the other contracted forms, on which the speakers disagree, is uncertain. It is possible that the relevant contractions are not stored, but result from a productive process contracting all combinations containing lax vowels, and that the corresponding full forms are not stored, either. In other words, contraction indicates that *heb ik*, *dat ik*, *weet ik*, *moet ik*, *had ik*, *laat ik*, and *deed ik* are probably stored in the lexicon, while it provides no information on other word-combination types.

In conclusion, frequency of occurrence and contraction do not give a decisive answer on exactly which verb form + *ik* combination types are stored in the lexicon. Since none of the verb form + *ik* combinations have lexical variants apart from the contracted ones, or any other properties which indicate that they are lexically stored, this implies that, on the basis of the available data, the category of *ik*-combinations cannot be divided into a group of types that are stored in the lexicon and a group of types that are not. Consequently, at present *Hypothesis III* cannot be tested by comparing these two groups.

Finally, *Hypothesis III* can be tested under the assumption that word-combination types with predictable properties are more likely to be retrieved as single units from the lexicon only if they are of an exceptionally high frequency of occurrence. If this assumption is correct, there is some evidence for *Hypothesis III*. It is a fact that the combination types *heb ik* /hɛb ɪk/ 'have I' and *dat ik* /dɑt ɪk/ 'that I', which are much more frequent than all other combination types, behave exceptionally (see Table 9.10). The type *heb ik* differs from the other types in that it is more likely to be classified with a voiced stop (Fisher's exact test, $p < 0.05$ two-tailed). This can be explained under the assumption that *heb ik* is more likely to be retrieved as a single unit from the lexicon. The word *heb* underlyingly ends in a /b/, and all models of the lexical forms of word-combinations (see Hypothesis 6 in §7.6) consequently assume that the lexical form of *heb ik* is /hɛbɪk/. Tokens of *heb ik* which are retrieved as units from the lexicon are therefore predicted to be realized with a voiced intervocalic stop. Hence, if *heb ik* is more likely to be retrieved as a single unit than other word-combination types, it is expected to be more likely to be realized with a voiced stop.

The combination type *dat ik* was less likely to be classified with a voiced intervocalic stop than all other combination types, including or excluding *heb ik* (Fisher's exact test, $p < 0.05$ two-tailed). In addition, the type was often realized as a monosyllable in our corpus (see Table 9.11), whereas this was rarely the case with the other *ik*-combinations. These data can be explained under the assumption that *dat ik* is stored as /dɑk/ and /dɑtɪk/ in the lexicon. The lexical form /dɑtɪk/ explains the high number of realizations classified with voiceless stops: it ensures that these realizations not only have a chance to surface when *dat* and *ik* are separately retrieved from the lexicon, but also when *dat ik* is retrieved as a single unit. The lexical form /dɑk/ explains why, although *dat ik* is stored as /dɑtɪk/, i.e. with a voiceless intervocalic stop, the proportion of tokens with voiceless stops for *dat ik* is not as high as the proportion of tokens with voiced stops for *heb ik*, which is stored with a voiced stop: it ensures that when *dat ik* is retrieved as a single unit from the lexicon, it is not always retrieved as /dɑtɪk/, but sometimes as [dɑk]. The assumption that *dat ik* has the lexical form /dɑtɪk/ implies that the intervocalic stop has the same [voice]-specification in the lexical form of *dat ik* as it has in the underlying form of *dat* (/dɑt/). These findings therefore support Booij's (1985) assumption that segments generally have the same specifications in the lexical form of a word-combination as in the underlying forms of the parts of the combination (§2.2.2).

Table 9.11 Absolute and relative (%) frequencies of occurrence of dat ik *in the corpus, broken down by three types of realization.*

Type of realization	Frequency of occurrence
[dɑk]	29 (32%)
with a voiced intervocalic stop	27 (29%)
with a voiceless intervocalic stop	36 (39%)

In conclusion, the combinations with *ik* provide evidence in favour of *Hypothesis III*, since this hypothesis can explain the exceptional behaviour of the highly frequent combinations *heb ik* and *dat ik*. The classifications of the intervocalic stop of *dat ik* suggest that the [voice]-specification of a word-final obstruent is the same in the underlying form of a word as in the lexical form of a string containing that word.

9.5.3 Obstruents in combinations ending in *het*

In addition to the combinations ending in *ik*, those ending in *het* may provide confirmation for Hypothesis III. Since none of these combinations seem to have unpredictable properties, and since we do not know the frequency threshold for storage (§9.5.2), it is impossible to determine which of them may have lexical representations. This means that the validity of *Hypothesis III* can only be investigated by examining the relation between the more or less continuously distributed variables "proportion of voiced stops" and "frequency of retrieval", and by investigating the realization of exceptionally highly frequent combination types (§9.5.2).

Table G.2 in Appendix G lists the frequency of occurrence of each combination type of verb form + *het*. The number of voiced and voiceless stops in these types as a function of the frequency of occurrence of the types can be found in Table 9.12.

Table 9.12 Absolute and relative (%) frequencies of occurrence of voiced and voiceless stops in combination types with het, *broken down by the frequency of occurrence of the combination types in the present corpus and the pilot study of the Eindhoven corpus of spoken Dutch.*

Frequency of occurrence	Instances of the combinations		Numbers of stops classified as			
			voiced		voiceless	
70	*heb het*	'have it'	6	(67%)	3	(33%)
59	*weet het*	'know(s) it'	12	(75%)	4	(25%)
36	*had het*	'had it'	1	(14%)	6	(86%)
10-15	*doet het* *deed het*	'does it', 'did it'	10	(71%)	4	(29%)
1-9	*snap het* *ziet het*	'understand it', 'sees it'	6	(67%)	3	(33%)

Under the assumption that the frequency with which a combination type is retrieved from the lexicon is proportional to its frequency of occurrence, *Hypothesis III* can be tested by means of Kendall's test with the more or less continuously distributed variables "proportion of voiced stops" and "frequency of occurrence". This test should be run separately on the combinations with underlyingly voiced and voiceless word-final stops, since the realization of *dat ik*, discussed in section 9.5.2, suggests that stops have the same specifications in the underlying forms of words as in the lexical forms of combinations containing these words. Therefore, word-final stops that are voiced in the underlying form of a word are expected to be more likely to be realized as voiced than stops that are underlyingly voiceless in the underlying form of a word when they are in onset positions in word-combinations with lexical representations. Kendall's test does not reveal any relation between the proportion of tokens with voiced stops and the frequency of a combination type either for the types with underlyingly voiced ($S = 3$, number of pairs = 4, $p > 0.05$, one-tailed) or the types with underlyingly voiceless stops ($S = 0$, number of pairs = 10, $p > 0.05$, one-tailed), and therefore provides no evidence for *Hypothesis III*.

Apparently, if the data present some evidence for *Hypothesis III*, this may only appear from an inspection of the behaviour of the exceptionally highly frequent combination types. Such an inspection, however, does not provide evidence for the hypothesis either, as there are no verb + *het* types which are of an exceptionally high frequency of occurrence, and those with a relatively high frequency do not behave differently from those with a relatively low frequency (see Table 9.12).

In conclusion, the data with respect to the verb + *het* combinations do not provide evidence for *Hypothesis III*. This does of course not imply that the hypothesis is incorrect. It is possible that no evidence has been found because all combination types are retrieved as single units from the lexicon equally often, or because the sample size is simply too small to reveal differences between the types.

9.5.4 Obstruents in combinations ending in *er* and in *met een*

The combination types containing *er* and the combination type *met een* are the last ones that are investigated for the evaluation of *Hypothesis III*. These combination types possess no known unpredictable properties. Therefore, if some, or all, of them are stored in the lexicon, this must be because they are highly frequent. Since we do not know the threshold frequency for storage, we cannot determine which types have lexical representations, and which ones have not. The validity of *Hypothesis III* can again only be investigated by examining the relation between the non-categorical variables "proportion of voiced stops" and "frequency of retrieval", and by investigating the realization of exceptionally highly frequent combination types (§9.5.2).

Table G.3 shows the numbers of voiced and voiceless intervocalic stops as a function of the frequencies of occurrence of the relevant combination types. Table 9.13 presents a summary of these data plus data on the combination types *dat er* and *met een.*

Table 9.13 Absolute and relative (%) frequencies of occurrences of voiced and voiceless stops in combination types with er, *and in* met een, *broken down by the frequency of occurrence of the combination type in the present corpus and the pilot study of the Eindhoven corpus of spoken Dutch.*

Frequency of occurrence	Instances of the combinations		Numbers of stop classified as			
			voiced		voiceless	
137	*dat er*	'that there'	0	(0%)	33	(100%)
109	*met een*	'with a'	14	(33%)	28	(67%)
45	*heb er*	'have there'	2	(18%)	9	(82%)
27	*gaat er*	'goes there'	1	(14%)	6	(86%)
25	*zit er*	'sit there'	3	(33%)	6	(67%)
15-20	*staat er* *had er*	'stands there', 'had there'	1	(20%)	4	(80%)
6-10	*ziet er* *zat er*	'sees there', 'sat there'	0	(0%)	8	(100%)
0-5	*koop er* *zet er*	'buy there' 'put there'	0	(0%)	9	(100%)

Running Kendall's rank-correlation test on the data in Table G.3 plus the data on *dat er* and *met een* again provides no support for *Hypothesis III*: the relation between the proportion of tokens with voiced stops and the frequency of occurrence of the combination type is statistically significant either for the combination types with underlyingly voiced stops ($S = 1$, number of pairs = 2, $p > 0.05$ one-tailed) or for the types with underlyingly voiceless stops ($S = 30$, number of pairs = 13, $p > 0.05$ one-tailed).

Nevertheless, the data present some support for Hypothesis III, since there is a statistically significant difference in the proportion of voiced stops between the combination types *dat er* and *met een* (Fisher's exact test, $p < 0.05$ two-tailed). Under the Complete Neutralization Hypothesis, the most plausible explanation for the difference is that, while all word-final stops of *met een* are unspecified for [voice], some word-final stops preceding *er* are specified as [−voice]. Stops preceding *er* are specified as such if they happen to precede the lexical form /dər/, instead of /ər/, of *er* (§8.2.2.2). They are then part of an alveolar geminate of which the second part belongs to a function word, and such geminates are likely to have the phonological specification [−voice] (cf. §7.5.3 and Chapter 10). Another plausible explanation for the difference between *dat er* and *met een* is that *dat er* has the lexical representation /dɑtər/, and is often retrieved as a single unit from the lexicon. There is some support for a lexical form of *dat er* with a voiceless stop in that the combination *dat ik* has a lexical form with a voiceless stop as well (§9.5.2).

Note that the combination *met een* cannot be assumed to have a lexical representation /mɛdən/. The problem is that it is unclear why *met een* would have a lexical representation with a voiced alveolar stop, whereas *met* has the underlying form /mɛt/, and listeners do not perceive [mɛdən] more often than [mɛtən].

9.5.5 Conclusions

The data provide some evidence for the hypothesis that the classification of intervocalic word-final stops is influenced by the frequency with which the combination types are retrieved as units from the lexicon. The fact is that the combination type *heb ik* is more likely, and *dat ik* is less likely to be classified with a voiced stop than the other combination types. This can be explained under the assumption that highly frequent word-combination types are often retrieved as single units from the lexicon, and that they have lexical forms in which the segments have the same specifications as in the underlying forms of the relevant words.

The data do not provide any stronger evidence for the hypothesis. One of the reasons for this may be that the data set is too small. Another reason may be that the data set contains too few combination types that are particularly likely or unlikely to be retrieved as single units from the lexicon.

9.6 Conclusions

This chapter evaluated the analysis proposed in Chapter 7 by evaluating *Hypotheses I* to *III* on the basis of the voiced/voiceless classifications of single intervocalic stops. The hypotheses that were tested are repeated below for convenience.

Hypothesis I

Obstruents which are lexically in coda positions are realized as voiced or voiceless independently of their underlying [voice]-specifications, and the phonological feature specifications of the adjacent segments, provided that the realization of these latter specifications does not interfere with the perception of voicing.

Hypothesis II

Obstruents which are lexically in coda position are more likely to be realized as voiced than onset obstruents that are phonologically voiceless, and less likely to be realized as voiced than onset obstruents that are phonologically voiced.

Hypothesis *III*

There is a systematic difference in realization between word-final obstruents before vowel-initial enclitics in word-combinations which are likely to be retrieved as single units from the lexicon and in word-combinations which are usually computed from their parts.

Hypothesis I is not falsified by the voiced/voiceless classifications of the stops, since these classifications do not suggest that the realization of a word-final stop as voiced or voiceless is related to its underlying [voice]-specification, or the roundness or laxness of the preceding vowel. The data set, on the contrary, strongly suggests that the phonological variables are of no influence.

Hypothesis II was confirmed by the classifications of the stops, since there appears to be a statistically significant difference between the classifications of word-final stops, and those of word-medial ones. As predicted, the word-medial stops are nearly always realized in accordance with their underlying [voice]-specifications, whereas the word-final ones are sometimes realized as voiced and sometimes as voiceless.

Finally, *Hypothesis III* could not be irrefutably confirmed by the classifications of the stops. The only supporting finding revealed by this study is that the two most frequent combinations behave exceptionally, and that their behaviour can be explained under *Hypothesis III*. It seems that a word-final intervocalic stop has the same [voice]-specification in the lexical representation of a word-combination as it has in the separate underlying form of the word of which it forms a part.

In conclusion, the classifications of intervocalic stops are in line with *Hypotheses I* to *III*. Consequently, they do not falsify the analysis developed in chapter 7.

10 Geminate alveolar stops

10.1 Introduction

This chapter will evaluate *Hypotheses I* and *IV* implied by the analysis proposed in Chapter 7 (§7.6) on the basis of obstruent clusters which are intervocalic and consist of a word-final alveolar stop and a word-initial /d/. Such clusters are contained, for instance, in the word-combinations *wat dan* /ʋɑt dɑn/ 'what then', and *had daar* /hɑd dar/ 'had there'. They are generally realized as single long stops.

The analysis claims that the first stop in these clusters is unspecified for [voice], as it is lexically in coda position. This stop therefore does not influence the realization of the geminate as voiced or voiceless. The realization of the geminate is completely determined by the phonological [voice]-specification of the second stop: it is realized as voiceless when this stop is phonologically [−voice], and as voiced when the stop is phonologically [+voice]. This second stop, which is underlyingly /d/, is phonologically [−voice] when it belongs to a word token that can be hypo-articulated to some extent, i.e., that is highly frequent and contributes little to the propositional content of the utterances.

The analysis therefore makes the following two predictions with respect to clusters consisting of a word-final alveolar stop and a word-initial /d/. These predictions are instances of *Hypotheses I* and *IV*, which were formulated in section 7.6.

Hypothesis I
The realization of the geminate is not influenced by the underlying [voice]-specification of its first part.

Hypothesis IV
The realization of the geminate is determined by the tendency of the /d/-initial word to be hypo-articulated, viz. by its frequency of occurrence, and its contribution to the propositional content of the utterance.

As mentioned in Chapter 8, the two hypotheses will be tested on the basis of all intervocalic /t d/ and /d d/ clusters from our corpus which were realized with single releases, and were unanimously classified as voiced or voiceless by ourselves and two other trained phoneticians. Appendix H lists the classifications of these clusters as a function of the type of word-combination in which they are contained.

We will evaluate the hypotheses by relating the classifications of the geminates as voiced or voiceless to the characteristics of the word-combinations containing the geminates. The acoustical properties of the geminates will not be taken into account, since these properties were not measured for the investigation of the validity of auditory classifications in section 8.3.5.2, and are therefore not available. Moreover, the acoustic properties of the geminates will not be considered because an investigation of acoustic properties contributes little to the evaluation of the hypotheses if it is based on a relatively small data set of geminates in very different contexts, as it would be in the present study (cf. §9.3.1.3).

For the evaluation of *Hypothesis I*, geminates starting with /d/s, i.e. /d d/-clusters, have to be compared to those starting with /t/s, i.e. /t d/-clusters. The second /d/ of nearly every /d d/-cluster belongs to a different type of word. This implies that word-combinations ending in different /d/-initial words have to be pooled for the testing of *Hypothesis I*. Which /d/-initial words have approximately the same influence on the realization of geminates, and therefore can be pooled, will become clear when *Hypothesis IV* is tested, which means that the testing of *Hypothesis I* depends on the testing of *Hypothesis IV*. This is why *Hypothesis IV* (§10.2) will be evaluated before *Hypothesis I* (§10.3).

In what follows the phrase "/d/-initial word realized with [t]" refers to a word in a word-combination that was perceived with a voiceless geminate. Similarly, a "/d/-initial word realized with [d]" refers to a word in a word-combination perceived with a voiced geminate.

10.2 The characteristics of the /d/-initial word

10.2.1 Overview of the data

It has generally been observed that the realization of an obstruent cluster ending in a word-initial /d/ is influenced by the type of word to which this /d/ belongs. If the /d/-initial word occurs in the list given in (1), the cluster is sometimes realized as voiced and sometimes as voiceless. Otherwise, the cluster is always realized as voiced (Leenen 1954; van Haeringen 1955; Demeulemeester 1962; Zonneveld 1982;

Gussenhoven 1989).

(1) Words of which initial /d/ can be realized as voiceless after obstruents

daar	/dar/	'there'
dan	/dɑn/	1. 'then' 2. 'than'
dat	/dɑt/	1. distal demonstrative determiner for neuter singular nouns 2. distal demonstrative pronoun for neuter singular nouns 3. relative pronoun for neuter singular nouns 4. conjunction
de	/də/	definite article for plural nouns and non-neuter singular nouns
d'r	/dər/	1. clitic form of *haar* 'her' 2. clitic form of *daar* 'there'
deze	/dezə/	1. proximal demonstrative determiner for plural nouns and non-neuter singular nouns 2. proximal demonstrative pronoun for plural nouns and non-neuter singular nouns
die	/di/	1. distal demonstrative determiner for plural nouns and non-neuter singular nouns 2. distal demonstrative pronoun for plural nouns and non-neuter singular nouns 3. relative pronoun for plural nouns and non-neuter singular nouns 4. clitic form of *hij* 'he'
dit	/dɪt/	1. proximal demonstrative determiner for neuter singular nouns 2. proximal demonstrative pronoun for neuter singular nouns
dus	/dʏs/	'thus'

The validity of this list was checked against the classifications of the alveolar stop geminates in our data set. The data set was restricted to geminates starting with word-final /t/, since the geminates starting with /d/ have only few tokens in the data set. The two types of geminates were not pooled, since the analysis was meant to be independent of the adopted analysis of word-final obstruents. If the two types of geminates are pooled, the assumption must be that the [voice]-specification of the first obstruent does not influence the realization of the geminates.

As indicated in (1), the tokens of some /d/-initial words may fulfill several functions. For instance, tokens of *dan* can function as an adverb ("then"), and a comparative conjunction ("than"). In general, tokens fulfilling different functions will be considered to belong to different types of words. Hence, the tokens of *dan* will be assumed to belong to either the adverb *dan*, or the conjunction *dan*.

The present tense and past tense forms of irregular verbs will also be considered to represent different types of words. The reason is that these forms represent different lexical entries, and could therefore be realized with [t] with different frequencies.

Most geminates in the data set are acoustically long. We saw in section 7.3.4 that acoustically long stops tend to be perceived as voiceless, which implies that it is more difficult to realize long stops as voiced than as voiceless, and that a geminate that is phonologically [+voice] is more often realized unfaithfully to its phonological [voice]-specification than a geminate which is phonologically [−voice]. The frequency with which phonologically voiced geminates are realized as voiceless is unknown, and cannot be determined on the basis of our data set, since it is also unknown which intervocalic alveolar geminates must be phonologically voiced.

Since we do not know the frequency with which phonologically voiced geminates are realized as voiceless, it is also unknown how frequently a /d/-initial word must be realized with [t] before it can be identified with any certainty as a word in which the /d/ can be phonologically voiceless. We assume here as a working hypothesis that the word must have at least two tokens with [t] in the data set, and that these tokens should form at least 10% of the total number of tokens of the word in the set. There are no words which happen to fulfill only the first part of the condition, and not the second part. Consequently, the words that fulfill both parts of the condition, and therefore will be assumed to be words that may have phonologically voiceless /d/, can be defined as words which have at least two tokens with [t] in the data set.

Table 10.1 lists the words (morphemes) that, according to this working hypothesis, may have phonologically voiceless /d/s. In addition, it lists for each of these words the numbers of occurrences with [t] and [d]. The words are ordered by decreasing proportion of realizations with [t]. See (1) for the underlying forms and (exact) functions of the words.

Table 10.1 The absolute and relative (%) frequencies of occurrence of voiced and voiceless geminates, broken down by the type of /d/-intial word. Only those /d/-initial words are included which occur at least twice in combinations with voiceless geminates.

Type of /d/-initial word	Numbers of geminates classified as			
	voiced		voiceless	
da 'k < *dat ik* 'that I'	0	(0%)	2	(100%)
dan conjunction	0	(0%)	2	(100%)
die relative pronoun	0	(0%)	2	(100%)
die 'he'	1	(2%)	54	(98%)
dan adverb	2	(3%)	59	(97%)
d'r 'there'	9	(8%)	106	(92%)
dat determiner	1	(13%)	7	(93%)
de	11	(14%)	72	(86%)
die determiner	12	(18%)	54	(82%)
die demonstrative pronoun	1	(20%)	4	(80%)
daar	3	(20%)	12	(80%)
dat conjunction	8	(21%)	30	(79%)
dat demonstrative pronoun	10	(25%)	30	(75%)
dit	2	(40%)	3	(60%)
dus	13	(81%)	3	(19%)

Since geminates tend to be realized as single stops (§3.5), /di/ meaning "he" and /dər/ are almost impossible to distinguish from their lexical variants /i/ and /ər/ when they follow alveolar stops. This implies that some tokens of *die* and *d'r* in the data set are probably vowel-initial instead of /d/-initial. When analysing the data, this fact should be taken into account.

Table 10.2 shows the morphemes that have at most one token realized with [t], and therefore perhaps do not occur with phonologically voiceless /d/. Those with one token with [t] are positioned above the dotted line, while the others can be found under the line. The morphemes are ordered alphabetically.

Table 10.2 The absolute and relative (%) frequencies of occurrence of voiced and voiceless geminates, broken down by the type of /d/-initial morpheme. Only those /d/-initial morphemes are included which are found at most once in combinations with voiceless geminates.

Type of /d/-initial morpheme			Numbers of geminates classified as			
			voiced		voiceless	
d'r	/dər/	'her'	0	(0%)	1	(100%)
dag	/dɑx/	'day'	0	(0%)	1	(100%)
dat	/dɑt/	rel. pron. (see 1)	0	(0%)	1	(100%)
deed	/ded/	'did'	5	(83%)	1	(17%)
deel	/del/	'part'	2	(67%)	1	(33%)
deze	/dezə/	see (1)	6	(86%)	1	(14%)
ding	/dɪŋ/	'thing'	10	(91%)	1	(9%)
doen	/dun/	'do' pres. tense	44	(98%)	1	(2%)
duur	/dyr/	'last'	3	(75%)	1	(25%)
daarom	/darɔm/	'therefore'	2	(100%)	0	(0%)
dacht	/dɑxt/	'thought'	1	(100%)	0	(0%)
denk	/dɛŋk/	'think' pres. tense	12	(100%)	0	(0%)
deug	/døx/	'be good'	2	(100%)	0	(0%)
diamant	/dijamɑnt/	'diamond'	1	(100%)	0	(0%)
diner	/dine/	'diner'	1	(100%)	0	(0%)
doel	/dul/	'target'	1	(100%)	0	(0%)
donker	/dɔŋkər/	'dark'	1	(100%)	0	(0%)
door	/dor/	'through'/ 'by'	3	(100%)	0	(0%)
dons	/dɔns/	'down'	2	(100%)	0	(0%)
dubbel	/dʏbəl/	'double'	1	(100%)	0	(0%)
duits	/dœyts/	'German'	3	(100%)	0	(0%)
duur	/dyr/	'expensive'	1	(100%)	0	(0%)
durv	/dʏrv/	'dare'	1	(100%)	0	(0%)

All words listed in Table 10.1 occur in (1). This implies that previous analyses identify all /d/-initial words which occur at least twice with [t] in our data set as words which can appear with devoiced /d/. These analyses apparently support the working hypothesis adopted here that a word can have a phonologically voiceless /d/ only if two as well as 10% of its tokens were realized with [t] in our data set. Conversely, our analysis shows that list (1) is probably complete.

Since previous analyses were mainly based on accidental observations, they could only distinguish between the types of /d/-initial words that can and those that cannot be realized with [t]. The present study is based on quantitative data, and can therefore investigate the frequencies with which /d/-initial words are realized with [t]. Table 10.1 shows that there are great differences among the words in this respect. Those differences that are statistically significant are listed in (2). The initial /d/ of each word on the left side of a ">"-sign in (2) occurs relatively more often in voiceless geminates than the initial /d/ of each word on the right side (Fisher's exact test $p < 0.05$, two-tailed).

(2) Hierarchies of words based on the frequency with which their initial /d/ is realized as [t]

a. *die* 'he' > *de*, *die* determiner, *dat* conjunction/pronoun, *dit*, *dus*
b. *dan* adverb > *die* determiner, *dat* conjunction/pronoun, *dus*
c. *d'r* 'there' > *dat* pronoun, *dus*
d. *de*, *die* determiner, *daar* adverb, *dat* conjunction/pronoun/determiner > *dus*

Hypothesis IV states that the /d/-initial words which can be realized with [t] comprise those which can be hypo-articulated to some extent, viz. the words which are of a high frequency of occurrence or contribute little to the propositional content of the utterance. We will test this hypothesis in sections 10.2.2 and 10.2.3. Section 10.2.2 will investigate the influence of the frequency of a word on the realization of its /d/, while section 10.2.3 will investigate the influence of its contribution to the propositional content.

Alternative explanations for the data could be that the realization of a word-initial /d/ is influenced by its historical origin, by the prosodic independency of the word, and by the presence of accent. Section 10.2.3 will discuss these possibilities, and argue that they are not necessarily relevant for an analysis of word-initial /d/ .

10.2.2 Frequency of occurrence

For the evaluation of the hypothesis that the frequency of occurrence of a word influences the realization of its /d/, we use the numbers of occurrence in our corpus and in the pilot study of the Eindhoven corpus to estimate the frequency of a word in spoken Dutch. We explained the reasons for using two corpora, and for using the Eindhoven corpus in particular, in section 9.5.1.

Table 10.3 lists the numbers of occurrence in the two corpora of the /d/-initial morphemes that probably can occur with a phonologically voiceless /d/, while Table 10.4 lists the numbers of occurrences of the other morphemes. The morphemes are ordered as in the corresponding Tables 10.1 and 10.2. The tokens of function words with identical appearance but with different functions (see 1), i.e. homonyms, are pooled for two reasons. The first one is that the word tokens in our corpus have not been tagged, which makes it impossible to distinguish homonyms and we do not have direct access to information on their frequency in this corpus. The second one is that Hierarchies (2) suggest that homonyms generally differ little in their realization. Only the tokens of *die* are split up into two groups. The reason is that tokens of *die* meaning "he" are more likely to occur with [t] than other tokens of *die* (see 2a). This difference is possibly related to the obligatory clitic status of *die* meaning "he" (see also §10.2.3).

Table 10.3 The numbers of occurrence of the words in our corpus and the pilot study of the Eindhoven corpus of spoken Dutch. The table is restricted to words which have at least two tokens in combinations with voiceless geminates.

/d/-initial word	Number of occurrence
da'k	minimally 29
die 'he'	526
dan	3444
d'r	2018
de	4979
die other function	4603
daar	1770
dat	7258
dit	323
dus	1804

Table 10.4 The numbers of occurrence of the morphemes in our corpus and the pilot study of Eindhoven corpus of spoken Dutch. The table is restricted to morphemes which have at most one token in combinations with voiceless geminates.

/d/-initial word	Number of occurrence
dag	141
deed	104
deel	82
deze	131
ding	304
doen	509
duur	55
daarom	54
dacht	205
denk	466
deug	4
diamant	4
diner	7
doel	17
donker	8
door	260
dons	12
dubbel	13
duits	39
duur	56
durv	2

The two tables show that nearly all words which can have phonologically voiceless /d/ are represented by more than 1000 tokens in the two corpora. In contrast, the other morphemes have frequencies of at most 509 tokens. A difference in frequency of occurrence of this size is probably sufficiently large to produce differences in realization. In general, the data appear to support the hypothesis that the possibility of an initial /d/ to be phonologically voiceless is related to the frequency of occurrence of the word.[1]

The three possible exceptions are *da'k*, *die* meaning "he", and *dit*, which are presented by 323 tokens or less in the two corpora, but are nevertheless regularly realized with [t]. The exact numbers of *da'k* in the two corpora is unknown, as it does not occur in the transcriptions of the pilot study of the Eindhoven corpus of spoken Dutch, which means that the form either does not occur in this corpus, or that it was transcribed as *dat ik*, the full form. It is possible that *da'k* and *die* meaning "he" were often realized with a [t], in spite of their low frequencies, because their realizations are influenced by the frequency of *dat*, which forms formally a part of *da'k*, and by the frequency of *die* with meanings other than "he", respectively. A further reason for the exceptional behaviour of *die* meaning "he" with respect to words of the same frequency could be the fact that this word is obligatorily enclitic (see also §10.2.3).

Whereas frequency of occurrence generally accounts for which /d/-initial words can have phonologically voiceless /d/, it does not account for the relative frequencies with which those that can have phonologically voiceless /d/ actually surface with [t]. This is evident from the results of Kendall's rank-correlation test when run only on the data in Tables 10.1 and 10.3 ($S = 15$, number of pairs $= 10$, $p > 0.05$. Homonyms, except those of *die*, have been pooled, since we do not have direct access to information on the frequencies of the different homonyms of a word in our corpus, see above). It is further supported by an investigation of the pairs of words which can have voiceless /d/ and for which the difference in the frequency with which they surface with [t] is statistically significant. Section 10.2.1 presented hierarchies showing which words are most likely to occur in combinations with voiceless geminates. These hierarchies are repeated in (3) for convenience, with homonyms being pooled.

(3) Hierarchies of words based on the frequency with their initial /d/ is realized as [t]

a. *die* 'he' > *de*, *die* other functions, *dat*, *dit*, *dus*
b. *dan* > *die* other functions, *dat*, *dus*
c. *d'r* > *dat*, *dus*
d. *de*, *die* other functions, *daar*, *dat* > *dus*

[1] This is also evident from Kendall's rank-correlation test over all data in Tables 10.1 to 10.4 ($S = 212$, number of pairs $= 31$, $p < 0.05$, homonyms have been pooled).

The words to the left of a ">"-sign are not always more frequent than the words to the right of that sign. The word *dan* (see 3b), for instance, is not represented by more tokens than *dat* and *die* in the two corpora (see Table 10.3).

Splitting up the homonyms probably does not make frequency of occurrence a better predictor. This is evident, for instance, for *dan* and *dat*, as the adverb *dan* and the demonstrative pronoun *dat* occur 1961 and 2513 times respectively in the Eindhoven corpus, although the adverb *dan* is realized with [t] more often than the demonstrative pronoun *dat* (see Hierarchy 2b).

In conclusion, the data show that the morphemes which can have a phonologically voiceless /d/ are generally more frequent than morphemes which cannot. Frequency of occurrence does not explain the relative frequencies with which /d/-initial words which may have phonologically voiceless /d/ actually surface with [t].

10.2.3 Contribution to the propositional content

Hypothesis IV states that the word tokens which are realized with [t] are not only highly frequent, but also not highly relevant to the propositional contents of the utterances in which they are contained. There is evidence for this part of the hypothesis, since the words which can be realized with [t] (see Table 10.1) are all principally relevant to the grammatical form of the sentence, and do not refer to a non-linguistic reality. They therefore belong to the group of "function words" as defined by van Wijk & Kempen (1980). The initial /d/ of content words is rarely realized as [t].

It is difficult to determine whether function words which are more often realized with [t] are generally less relevant to the propositional content on objective grounds. The problem is that there are no principled criteria to determine the relative relevance of function words. This means that the relation between the realization of a /d/-initial word and its relevance to the propositional content of the utterance can only be investigated if arguments are accepted which are partly based on intuitions.

There are four arguments of this kind for the existence of a relation. The first argument is that the only two tokens of the adverb *dan* in the data set which were realized with [d] refer to particular, fixed, points in time (see 4), whereas none of its tokens realized with [t] do. The latter tokens mean, for instance, "after that" (see example 5a), "in that situation" (example 5b), or are part of fixed expressions, such as *wat dan ook* /ʋɑt dɑn ok/ 'whatever'. If the tokens that refer to particular, fixed, points in time are removed from the sentence, the propositional content of these sentences becomes ambiguous, whereas this is not the case if the other tokens of *dan* are removed. Tokens of *dan* which refer to points in time are apparently more relevant to the propositional content of an utterance than the other tokens. The fact that it is precisely these tokens that occur in combinations with voiced geminates, then, supports the claim that the contribution of a /d/-initial word to the proposition influences the realization of its /d/. The evidence is, however, not strong, since it is based on only two tokens of *dan* realized with [d].

(4) The word *dan* (underlined) realized with [d]. Its "antecedents" are printed in small capitals. The sentences have been translated freely.

a. Subject G: *IN DE 19E EEUW was het een (..) heel beruchte stad*
IN THE 19TH CENTURY it was a (...) very notorious city

met veel (...) stokerijen, (..) mooie molens langs
with many (...) distilleries, (...) beautiful mills along

de grachten, de vesten zoals dat dan heette.[2]
the canals, the moats as they were then called.

b. Subject M: *Ik moest 3 JANUARI (...).*
I had to (...) on 3 JANUARY (...).

Dus ik zat ook al met het idee van
So, I had already the idea at the back of my mind of

Ik ben benieuwd of het dan sneeuwt
I am curious whether it will then be snowing

en hagelt om 9 uur 's ochtends.
and hailing at 9 o'clock in the morning.

(5) The word *dan* (underlined) realized with [t]. The sentences have been translated freely.

a. Subject H: *Je hebt de kamers en de boezems van een hart*
There are the ventricles and the auricles of a heart

en normaal gesproken komt de prikkel uit
and normally the stimulus leaves

de boezems en gaat dan naar de kamers
the auricles and then goes to the ventricles

---

2 The symbol "(...)" indicates hesitations in this example.

b. Subject M: *Of lever je ze ook los?*
Or do you also provide them separately?

En zo ja, wat moet dat dan kosten?
And if so, what would I then have to pay?

The second argument is based on the behaviour of *dus*. The three tokens of *dus* which were realized with [t] convey pragmatic information: they indicate that the speaker knows that what he says is no news to the listener (see 6). In contrast, most tokens of *dus* realized with [d] announce the repetition of relevant information (example 7a), or introduce logical conclusions of the preceding proposition (example 7b). Since repetitions of relevant information generally introduce new topics in the conversation, the announcements of repetitions are highly relevant to the propositional content of an utterance. They are probably more important for communication than the indication of pragmatic relations. Similarly, the introduction of a logical conclusion is probably also more important. The data therefore seem to support the hypothesis that initial /d/ is realized as [t] more often in words that contribute little to the proposition. Again, the evidence is not strong, particularly because the data set contains only three tokens of *dus* realized with [t].

(6) The word *dus* (underlined) realized with [t]. The sentences have been translated freely.

a. ME: *Er zijn nu van die ozonbaden.*
Nowadays there are those swimming pools with ozone.

Subject B: *Ja, waar?*
Yes, where?

ME: *In Amstelveen.*
In Amstelveen.

Subject B: *Ja, want dat Keerpunt, daar ging ik vroeger zwemmen.*
Yes, because that Keerpunt, I used to swim there.

Dat bestaat dus helemaal niet meer.
That [discourse marker] exists no longer.

b. Subject L: *maar bij supermarkten lag verkeerd vlees*
but there was wrong meat at the supermarkets,

over de datum heen. Hé, het is bijna,
beyond the sell-by date. Hey, it is almost,

ja nou ja je weet _dus_ *dat*
yes well yes you know [discourse marker] that

ik behalve een vriend van jou ook nog verkoper ben.
I am not only a friend of yours but also a salesman.

c. Subject O: *dat we (...) allerlei dingen hebben georganiseerd*
that we have organized all kinds of things (...)

van zeg maar kerstdiner,
from, say, Christmas dinners,

wat _dus_ *elk jaar gehouden wordt, tot en met ...*
which [discourse marker] is held every year, until...

(7) The word *dus* (underlined) realized with [d]. The sentences have been translated freely.

a. Subject E: *Ik werk nu nog alleen nog op maandag en*
I work only on Mondays these days and

vrijdagavond (...). Nou dan is het zo. Ik begin dan...
Friday evenings (...). Well, it is then like this. I then start...

Ik doe dat _dus_ *maandag en vrijdag. En dan*
So that's what I do on Mondays and Fridays. And then

begin ik om 2 uur 's middags. Dus in principe tot
I start at 2 o'clock in the afternoon. So in principle until

10 uur, half 11 's avonds.
10, half past 10 in the evening.

Dus, ja, het zijn gewoon dagen van 8 uur.
So, yes, these are regular days of 8 hours.

b. Subject B: *100 gulden was ook de prijs*
100 guilders was also the price

die ik voor de huidige A heb betaald dit jaar.
that I paid for the present A this year.

Je gaat _dus_ *niet omhoog (met je prijs).*
So you are not raising (your price).

The third argument concerns the determiners *de*, *die*, *dat*, *dit*, and *deze*. Here, too, there appears to be a relation between the realization of a /d/ and the propositional relevance of the word. The determiners *de*, *die* and *dat* were realized with [t] in an average of 85% of cases and the determiners *dit* and *deze* in only an average of 33% of cases (see Table 10.5). The difference is statistically significant (Fisher's exact test, $p < 0.05$ two tailed).

Table 10.5 The absolute and relative (%) frequencies of occurrence of voiced and voiceless geminates in combinations ending in a determiner, broken down by the type of determiner.

Type of determiner	Numbers of geminates classified as			
	voiced		voiceless	
de, die, dat	24	(15%)	133	(85%)
deze, dit	8	(66%)	4	(33%)

The forms *de*, *die*, and *dat* rarely contribute much to the propositional content of an utterance, since *de* invariably, and *die* and *dat* almost invariably, simply indicate that the referent of the following noun is assumed to be known, or finds its referent in the directly preceding discourse, which may be obvious (see examples 8).The forms *dit* and *deze* are more important. They precede and signal nouns which are important entities in the discourse, but of which the referents may not be obvious, as these referents may not have been described in the directly preceding discourse but before, or must be chosen among a number of possible referents. In other words, *dit* and *deze* signal to the listener that it is important to recover the referents of the following nouns, although it may take a special effort (Kirsner 1987 et al.; Kirsner & van Heuven 1988; see example 9 from the corpus). This means that these determiners fulfil an important function. Hence, the fact that it is especially these determiners that are generally realized with [d] supports the claim that the relevance of a word influences the realization of its /d/.

(8) Determiners (underlined) realized with [t] . The sentences have been translated freely.

a. Subject C: *Ik moet gewoon goede slaapzakken hebben in de ...*
I just need good sleeping bags in the ...

Subject D: *Ja, dat begrijp ik, dat begrijp ik.*
Yes, I understand, I understand.

Subject C: *Een goede voorraad goede slaapzakken.*
A good supply of good sleeping bags.

Ik zit alleen een beetje met <u>die</u> prijs.
I just have a problem with <u>that</u> price.

b. Subject G: *Je hebt zijn artikel wel eens gelezen over die kat van*
Have you ever read his article about that cat of

Schrödinger? (...). Dat ie iets met een experiment
Schrödinger? (...). That he something with an experiment

waarbij dus de tijd achterstevoren ging, geloof ik.
where the time went back, I believe.

Je kent dat niet? Nou hij leest altijd
You don't know it? Well, he always reads

<u>dat</u> soort boeken die (...) beetje science fiction-achtig zijn.
<u>that</u> type of books which (...) are a bit science fiction-like.

(9) Determiner (underlined) realized with [d]. The sentences have been translated freely.

a. Subject M: *Ja, ja en een andere fabrikant,*
Yes, yes, and another manufacturer,

daar heb je dan ook geen zin in?
you then don't feel like that either?

Subject N: *Nou ik heb, nee, ik heb een goede relatie*
Well, I have, no, I have a good relationship

met <u>deze</u> fabrikant.
with <u>this</u> manufacturer.

The final is based on the exceptionally high percentages of tokens of *die* meaning "he" and *d'r* realized with [t]. These words are clitics (Booij 1995:167), and therefore cannot bear stress, and do not provide new information. This implies that the observation that they have high percentages of tokens with [t] supports the analysis that /d/-initial words are more likely to be realized with [t] if they contribute little to the propositional content.

In conclusion, the data suggest that /d/-initial words which are often realized with [t] tend to contribute little to the propositional content of the utterances, which is in accordance with *Hypothesis IV*. *Hypothesis IV* cannot be tested definitively on the present data set because this data set is relatively small, and there are no principled criteria to determine the contributions of function words to the propositional contents of the utterances.

10.2.4 Three more characteristics of the /d/-initial words

The data in sections 10.2.2 and 10.2.3 suggest that /d/-initial words which may be realized with [t] are highly frequent or contribute little to the propositional contents of the utterances. These data therefore seem to support the hypothesis that /d/-initial words can be realized with [t] if they can be hypo-articulated to some extent.

It could also be argued, however, that the fact that some /d/-initial words can be realized with [t], whereas others cannot, is partly due to another difference between the two types of words. Two possibilities are mentioned in the literature.

First, Zonneveld (1983) argues that the difference in possible realization is due to a difference in prosodic status: it would result from the fact that some /d/-initial words can be incorporated into the preceding prosodic word, in which they are realized as [t] if preceded by obstruents, whereas others cannot (§3.4.5). Zonneveld's proposal fails to account for all the data. Some [t]s in the data set are acoustically relatively long, and therefore must be geminates in the output of phonology. Since geminates are banned from occurring within prosodic words (§3.5), these [t]s must span a prosodic word boundary. This implies that the underlying /d/ of the clusters to which these [t]s correspond belong to /d/-initial words that form prosodic words of their own. It seems, then, that /d/-initial words can be realized with [t] even if not incorporated into the preceding prosodic word. The problem of which words can and which ones cannot be realized with [t] cannot be accounted for by assuming a difference in prosodic structure.

Second, it has been observed that initial /d/s which can be realized as [t]s were underlyingly /θ/ in West-Germanic (van Ginneken 1935: 101, 102; Leenen 1954, 1955; Gussenhoven & Bremmer 1983: 66, 67). Gussenhoven & Bremmer (1983) suggested the following explanation for this fact. The /θ/-initial words which normally bore stress were probably consistently realized with [θ] in West-Germanic, whereas those which were generally unstressed were sometimes realized with [θ] and sometimes with [ð]. The unstressed ones consequently constituted a separate group, and could evolve differently. They evolved into words which are sometimes realized with [t] and sometimes with [d] in Modern Dutch, whereas the stressed /θ/-initial words and the /d/-initial ones evolved into words which are always realized with [d]. Only the initial /d/ of words that are unstressed and were /θ/-initial in West-Germanic can therefore be realized as voiceless in Modern Dutch. The reason why the words in (1) can be realized with [t], whereas, for instance, the word /dor/ 'through'/'by' cannot would then be related to their difference in origin.

The assumption that there is a relation between the modern realization of a /d/ and its historical origin is plausible. There are several other phenomena that can be accounted for by referring to a previous language stage. The pronunciation of the second *z* in *zevenenzeventig* 'seventy-seven' as an [s] in Modern Dutch is a case in point. This /z/ is often realized as [s] ([zevənəsevəntəx]) probably because it was preceded by a /t/ in Middle-Dutch, and consequently underwent Progressive Voice Assimilation (Brink 1970: 175, 176; van Reenen & Wattel 1992).

An appeal to data from previous language stages is, however, unnecessary for a proper account of the realization of initial /d/ as [t] in Modern Dutch. All data can be explained under the assumption that only those words can be realized with [t] that can be hypo-articulated to some extent, i.e., that are highly frequent or contribute little to the propositional content of the utterance. A good example is the word *door*, which is rarely realized with [t]. This word has a frequency of only 260 tokens in our corpus and the Eindhoven corpus, and is generally more relevant to the propositional content of the utterance than definite articles and pronouns, since it is a preposition and therefore indicates the grammatical function of the following Noun Phrase. Given its non-high frequency of occurrence and its relevance to the propositional content of the utterance, the word is not expected to be clearly hypo-articulated, i.e. realized with [t]. If we accept the assumption that words can be realized with [t] only if they can be hypo-articulated to some extent, there is no need to appeal to historical data.

There is also a problem with an account based on historical data: it offers no explanation, for instance, for the fact that *dan* is more often realized with [t] than *dat*, *die*, *dit*, and *dus* (see 2). Hence, it has to be supplemented by an assumption which can explain the frequencies with which words are realized with [t]. In contrast, the assumption that the realization of an initial /d/ as [t] is an instance of phonologized hypo-articulation can explain all the findings on its own.

There is a third difference between the words that can and those that cannot be realized with [t], since generally only words without accent are realized with [t]. This difference follows naturally under the assumption that the realization of a word-initial /d/ as [t] is an instance of phonologized hypo-articulation. The reason is that the accent level of a word is determined by, and therefore derived from, its relevance to the propositional content of the utterance. Note that the factor "accent" cannot replace the factor "contribution to the proposition", because "accent" cannot distinguish between the various functions of function words, and therefore cannot explain the observations discussed in section 10.2.3. The assumption that accent plays an independent role in an analysis of word-initial /d/ is unnecessary.

In conclusion, there are no reasons to assume that analyses of word-initial /d/ must refer to the historical origin of the /d/, or the prosodic dependency or accentedness of the word. Analyses which assume only these factors can explain fewer facts than an analysis which adopts the idea of phonologized hypo-articulation. This conclusion is in line with *Hypothesis IV*.

10.3 The [voice]-specification of the word-final stop

The second hypothesis that will be tested on the basis of the alveolar geminates is *Hypothesis I*. This hypothesis states that the classification of a geminate as voiced or voiceless is not related to the underlying [voice]-specification of its word-final stop, since this stop is unspecified for [voice].

For the evaluation of the hypothesis, geminates starting with word-final /t/ must be compared to geminates starting with word-final /d/ with respect to their classifications as voiced or voiceless. Ideally, the investigation should compare geminates of which the second part, i.e. the word-initial /d/, belongs to the same word type. If differences are found between the two types of geminates, these differences then cannot be due to the influence of the word-initial /d/, and must consequently be due to the influence of the word-final stop.

Unfortunately, in the present study the evaluation cannot be based on combinations ending in the same words, since the data set contains too few examples. We pooled the combinations ending in *dat* and *die*, with the exception of combinations ending in tokens of *die* meaning "he", and tested the hypothesis on the basis of just these combinations. These combinations were chosen because they seem to form a sufficiently large data set, represent voiced and voiceless geminates, and their final words probably have approximately the same influence on the realization of geminates and consequently can be pooled (see hierarchy 2 in §10.2.1).

We assume that words which are written with *d* have underlying /d/, and that words written with *t* have underlying /t/. For instance, we assume that the word *goed* 'good' ends underlyingly in /d/, and that the word *niet* 'not' ends underlyingly in /t/. The assumption is probably only incorrect for words ending in *d* which are stored in the lexicon separately from related forms. On the basis of the observation that their word-final stops are generally realized as voiced, these words could be argued to end in /t/, although they are realized with [d] when incorporated into other words or strings of words. The word *altijd* [ɑltɛit] 'always', for instance, could be argued to have an underlying word-final /t/, although related to word *tijden* [tɛidən] 'times', which is realized with [d], because it is a fossilized combination (/ɑl+tɛid/ 'all time'), and is nearly always realized as [ɑltɛit]. We did not attempt a classification of words ending in *d* as words that probably end in /t/ or /d/ along these lines, because it seems unlikely that such a classification could ever be based on anything but intuitive grounds, and would therefore be of only limited value for an objective analysis. We will assume here that the proportion of potentially problematic words is too small to influence the results of the evaluation of *Hypothesis I*.

Table 10.6 shows the frequencies of occurrence of voiced and voiceless geminates in combinations ending in *dat* or *die* with *die* meaning something other than "he". The numbers are broken down by the underlying [voice]-specification of the final stop of the first word. Since *Hypothesis I* states that these [voice]-specifications do not influence the classifications, it is falsified if an effect is attested with an associated p-value equal or smaller than 0.2 (§8.4). The data in the table do not present such an effect: they show no statistically significant difference between the proportions of voiced geminates among combinations with underlyingly voiced and voiceless stops (Fisher's exact test, $p > 0.2$ two-tailed).

Table 10.6 The absolute and relative (%) frequencies of occurrence of voiced and voiceless geminates in combinations ending in dat *and* die, *except* die *meaning "he". The geminates are broken down by the underlying [voice]-specification of the word-final stop.*

Underlying specification	Numbers of geminates classified as			
	voiced		voiceless	
[+voice]	8	(27%)	22	(73%)
[−voice]	32	(20%)	128	(80%)

If the proportions of voiced geminates differ, the maximal difference is 0.20 with a certainty of 95% (formula 2 in §8.4). A difference of such a size is probably sufficiently small to be due to factors which influence the realizations of the geminates without changing their [voice]-specifications.

If there is a difference, and the geminates starting with /t/ are more often realized as voiced than geminates starting with /d/ in the population, this can be due to the exceptional behaviour of word-combinations starting with *niet* 'not'. The word *niet* has two lexical representations: /nit/ and /ni/ (§6.2.2.4). Word-initial /d/ following /ni/ is not preceded by an obstruent, and is therefore phonologically voiced, and realized as voiced. Hence, combinations starting with *niet* are predicted to be more often realized with voiced stops than other word-combinations. Table 10.7 shows that this prediction is accurate. The combinations *niet die*, with *die* meaning something other than "he", and *niet dat* were more often realized with voiced stops in our corpus than the other combinations ending in *die* meaning something other than "he", or *dat* (Fisher's exact test, $p < 0.05$ one-tailed). Recall that it is difficult to perceive whether a word-combination contains a single stop or a geminate, since a sequence of two obstruents of the same place and manner of articulation is generally not realized as two separate obstruents, but as a single (long) one.

Table 10.7 Absolute and relative (%) frequencies of occurrence of voiced and voiceless stops in combinations ending in die *meaning something other than "he" or* dat *and starting with* niet *and in those starting with another type of word.*

Type of first word	Geminate classified as			
	voiced		voiceless	
niet	6	(40%)	9	(60%)
Other word	27	(16%)	141	(84%)

In conclusion, the data do not suggest that the realization of geminates spanning word-boundaries is influenced by the underlying [voice]-specifications of their first, word-final, stops. This finding is in accordance with *Hypothesis I*.

10.4 Conclusions

This chapter evaluated the analysis presented in Chapter 7 on the basis of geminates consisting of a word-final alveolar stop and a word-initial /d/. The hypotheses that were tested are the following.

Hypothesis I
The realization of the geminate is not influenced by the underlying [voice]-specification of its first part.

> *Hypothesis IV*
> The realization of the geminate is determined by the tendency of the /d/-initial word to be hypo-articulated, viz. by its frequency of occurrence, and its contribution to the propositional content of the utterance.

It appeared that the data confirm both hypothesis. *Hypothesis I* is supported because the data do not suggest that the realization of a geminate as voiced or voiceless is influenced by the underlying [voice]-specification of its first, word-final stop. *Hypothesis IV* is supported since the data indicate that nearly all voiceless geminates end in the /d/ of a highly frequent word, and that the general contribution of a /d/-initial word to the propositional content of the utterance is related to the relative frequency with which its /d/ is part of a voiceless geminate.

The hypotheses could not be conclusively tested, since it is unknown which morphemes can be pooled in an analysis of frequency effects, and what are the exact contributions of words to the propositional content of the utterances. A second reason why conclusive evidence could not be provided is that the data set is relatively small.

In summary, the behaviour of the geminates is in line with *Hypotheses I* and *IV*. The single stops discussed in Chapter 9 are in line with *Hypotheses I* to *III*. Therefore, all the data discussed in this study are in line with the analysis of the realization of obstruents as voiced or voiceless in Dutch that was proposed in Chapter 7. They allow for an analysis which considers coda obstruents as phonologically unspecified for [voice], and assumes that word-initial /d/s may be realized as voiceless in obstruent clusters if they belong to words that may be hypo-articulated.

Part VI

Final Remarks

11 Conclusions

11.1 Introduction

The preceding chapters have been dedicated to a study of segment reduction and obstruent voicing in casual Modern Standard Dutch which aims to present new data, and to shed light on the question whether the characteristics of casual Dutch are due to the speaker's natural tendency to reduce articulatory effort, and which other factors are relevant. We presented a rough survey of the contexts in which segments may be absent and vowels may be realized as schwa, and developed and tested a new analysis of the realization of obstruents as voiced or voiceless which does justice to data from fluent speech, and incorporates the assumption that the realization of coda obstruents and word-final obstruents as voiced or voiceless depends on which realization requires no additional articulatory effort. An additional focus of attention was the research method used, which was discussed in great detail.

This chapter will summarize the most important conclusions that were reached in each part of the book, and relate the results of the different parts. Parts I and II are not included in this summary, since they merely presented background information.

11.2 Part III: Type of data

Part III of this book motivated the choice for the type of data which formed the basis for the investigations. Chapter 4 started with the claim that studies on the realization of segments in casual speech cannot be based exclusively on linguistic intuitions, since such intuitions are not valid with respect to the non-lexical properties of words. This claim was supported by the results of a brief investigation of the intuitions of Dutch speakers on the realization of word-final intervocalic stops

(§4.2.2), which showed that speakers of the same variant of Dutch may differ widely in their intuitions.

The speakers whose intuitions were investigated included the 16 subjects who produced speech for the compilation of the corpus. Their realization of word-final intervocalic stops was investigated in Chapter 9, and the validity of these speakers' intuitions can therefore be directly studied in a comparison of their intuitions and their actual behaviour. Table 11.1 presents the intuitions of eleven of these subjects with respect to post-vocalic /d/s and /t/s before *ik* and *het*. Table 11.2 presents the number of such stops that were realized as voiced and voiceless in the corpus.

Table 11.1 The intuitions of 11 speakers on their realization of intervocalic word-final alveolar stops. The stops are broken down by underlying [voice]-specification and type of following function word.

Stop	Possible realization according to			
	Subject B	Subjects D, E, G, J, L, N	Subject K	Subjects I, M, O
/...Vd̲ ɪk/	[d]	[t]	[t] and [d]	[t] and [d]
/...Vt̲ ɪk/	[d]	[t]	[t]	[t] and [d]
/...Vd̲ ət/	[t] and [d]	[t]	[d]	[t]
/...Vt̲ ət/	[t] and [d]	[t]	[t] and [d]	[t]

Table 11.2 Absolute and relative (%) frequencies of occurrence of voiced and voiceless intervocalic stops in the data set, broken down by underlying [voice]-specification and type of following function word.

Stop	Numbers of stops classified as			
	voiced		voiceless	
/...Vd̲ ɪk/	31	(79%)	8	(21%)
/...Vt̲ ɪk/	93	(80%)	23	(20%)
/...Vd̲ ət/	7	(54%)	6	(46%)
/...Vt̲ ət/	21	(68%)	10	(32%)

The data from the corpus show that both underlying /t/ and /d/ can be realized as voiced and as voiceless before *ik* and *het*. This is not in line with the intuitions summarized in Table 11.1, demonstrating that these intuitions are not a reliable guide to actual behaviour. There is no support for the assumption that a subject's intuitions may agree with his own realizations even if they do not with the data from the corpus as a whole, since the data set does not show that the subjects differ in their realizations. The comparison therefore provides additional evidence for the hypothesis that many speakers do not have valid intuitions on the realization of intervocalic word-final stops.

The second part of Chapter 4 dealt with speech recordings. It was argued that they constitute the most important type of data for studies of casual speech, since all other types of data, such as linguistic intuitions, speech errors, and sound changes, cannot provide sufficient and valid evidence. Recorded stretches of speech constitute valuable data when their perceptible characteristics have been transcribed as strings of symbols. The transcription of some characteristics can partly be based on acoustic measurements, whereas the transcription of others can only be made by ear. Transcribing by ear is difficult, which means that auditory transcriptions are, ideally, only accepted if they have been unanimously arrived at by several independent transcribers. It was shown in Chapters 6 and 8 that this is actually an important requirement, since transcribers often disagree on the transcription of stretches of casual speech. For instance, two independent transcribers and myself disagreed on the presence/absence of the vowel following the [n] in *natuurlijk* /natyrlək/ in 58% of cases (§6.4.2), and on the classification of intervocalic stops as voiced or voiceless in 15% of cases (§8.3.3).

Chapter 4 ended with a discussion of the advantages and disadvantages of recordings of scripted speech, and recordings of unscripted speech, i.e. linguistic corpora. The most important advantage of recordings of scripted speech is that it is easy to ensure that they contain the realizations in which one is interested. Linguistic corpora have the advantage that offer the best chances of representing casual speech, and that the same corpus may form a basis for surveys of many different phenomena, including phenomena of which little is known. This is why corpora constitute a more suitable basis for the intended investigations, and why the present study was based on a corpus.

Chapter 5 described the corpus used, which was constructed especially for the purposes of the present study. This corpus consists of approximately 122,500 word tokens realized by 16 male speakers in 12 hours of conversation. The corpus was designed so that it would meet, as far as possible, 8 general requirements which follow from the purposes of this study (§5.1). These requirements are repeated in (1).

(1) Requirements for the corpus

1. The corpus should contain casual speech.
2. The corpus should represent Standard Dutch.
3. The subjects should speak the same variety of Standard Dutch.
4. The recordings that make up the corpus should be available.
5. The recordings that make up the corpus have to be free of background noise.
6. The corpus should not be very large.
7. The corpus should be sufficiently large for the purposes of the investigations.
8. Every speaker should be represented by a relatively large number of realizations.

The corpus meets Requirements (1-4) and (6) as well as possible. The other requirements were fulfilled to a lesser extent. First, the corpus does not completely meet Requirement (5), since it contains some background noise, even though it was tape-recorded in a soundproof room (§8.2.4). Second, it does not completely meet Requirement (7), since it proved to be insufficiently large to allow for the conclusive testing of all hypotheses incorporated by the analysis proposed in Chapter 7 (see Chapters 9 and 10). Finally, the corpus does not meet Requirement (8), since it contains too few realizations of intervocalic obstruents per speaker to allow for studies of interspeaker variability (§8.4).

Although the corpus is not perfect, it provides valuable data for linguistic analysis. This was shown in parts IV and V of the book.

11.3 Part IV: A rough survey of phoneme realizations

Chapter 6 of this book presented a rough survey of the contexts in which segments that are present in highly careful speech can be absent and vowels can be realized as schwa in casual Dutch.

The data allow for the following generalizations.

(2) Generalizations with respect to the absence of consonants.

a. [t] is primarily absent, when it is expected to
   - be in coda position and follow [s];
   - precede a bilabial stop;
   - form the final segment of *niet* /nit/ 'not';
   - form the final segment of a highly frequent verb-stem.

b. [r] is primarily absent, when it is expected to
   - be in coda position and follow [ə];
   - follow low vowels;
   - be part of the word *precies* /prəsis/ 'exactly'.

c. [n] is primarily absent, when it is expected to
   - be in coda position and follow [ə];
   - be in coda position and precede an obstruent.

d. [d] can may absent, when it is expected to follow [n] and precede [ə].

e. Other types of consonants may be absent in the acoustic forms of a number of words at least. They are:
   - the [h] in forms of *hebben* /hɛbən/ 'to have';
   - the [x] in *nog* /nɔx/ 'yet' and *toch* /tɔx/ 'nevertheless';
   - the [k] in forms of *denken* /dɛŋkən/ 'to think';
   - the [l] in *als* /ɑls/ 'if';
   - the [f] in *zelfs* /zɛlfs/ 'even' and *zelfde* /zɛlfdə/ 'same'.

(3) Generalizations with respect to the realization of underlyingly full vowels as schwa.

a. Phonological monophthongs (cf. §3.2) as well as the diphthong /ɛi/ may be realized as schwas.

b. Vowels may be realized as schwas in open and closed syllables, and in word-medial and word-final syllables.

(4) Generalizations with respect to absence of vowels.

a. The vowels [ɪ] and [ə] are generally absent when they are expected to be adjacent to a vowel that belongs to the same prosodic word.
b. Schwas are optionally absent when they are expected to follow an obstruent and precede a liquid and an unstressed vowel within the same prosodic word.
c. All vowels are optionally absent when they are expected to be adjacent to continuants, especially fricatives.

It could be argued that the perceptual absence of some types of segments in some contexts, such as coda [t] after [s], and coda [r] after [ə], might be exclusively due to the speaker's natural tendency to reduce articulatory effort. These segments are generally acoustically non-salient, belong to highly frequent items, and do not occupy positions in the word which are highly relevant for recognition, which implies that their absence does not block communication. In addition, these segments may be perceptually absent when they are co-articulated with the adjacent gestures or when the speaker reduces the sizes of the articulatory gestures. The absence of these segments seems to support the claim that the speaker's natural tendency to reduce articulatory effort plays an important role in casual speech.

The observation that particularly highly frequent items surface in reduced forms is probably not only due to the fact that these words are easily recognized by the listener. Other possible causes are that reduced forms of high frequency items may be stored in the lexicon, and have a higher probability to be incorporated into a non-systematic data set.

The data show that more than one segment may be absent in the acoustic forms of items. The attested maximally reduced forms indicate that all segments of unaccented items may be absent in such forms, except the initial and final ones, and the ones in the onsets and nuclei of stressed syllables. That is, only segments which are prominent in storage (§2.2.2) must be realized in highly reduced forms. The reduced forms of an item make up a multi-dimensional continuum between the full form of the item and its maximally reduced form.

11.4 Part V: The realization of obstruents as voiced or voiceless

Previous analyses of the realization of obstruents as voiced or voiceless are problematic, because they cannot account for the data which are discussed in studies by Meinsma (1958), Kaiser (1958), Demeulemeester (1962), Slis (1982, 1983), and others. Chapter 7 proposed an analysis which is in line with all data, and ties in with the assumption that the speaker's natural tendency to reduce articulatory effort plays an important part in casual speech.

The analysis states that obstruents which are in coda position at the lexical level of phonology are unspecified for [voice] in the output of phonology, and in the input of phonetics (Complete Neutralization Hypothesis). They are realized as voiced when a voiced realization is easier than a voiceless one, and as voiceless when a voiceless realization is easier, with easier being defined as requiring smaller and less exactly timed articulatory gestures. In contrast, obstruents which are not in coda position are specified for [voice] in phonology as well as in the input of phonetics, and are realized in accordance with their phonological [voice]-specifications.

In addition, the analysis states that all underlyingly voiced fricatives as well as the initial /d/s of some function words are realized as voiceless after obstruents, because the constraint which bans obstruents in clusters from being [+voice] dominates the members of the constraint family IDENTFEATURE(VOICE) on fricatives and the segments of the relevant function words. The low ranking of the constraint IDENTFEATURE(VOICE) on fricatives expresses the general weakness of the voiced/voiceless distinction for fricatives. The low ranking of the constraint IDENTFEATURE(VOICE) on the segments of the relevant function words is in line with the observation that function words may be hypo-articulated. They may be hypo-articulated since they are highly frequent and generally contribute little to the propositional content of the utterances.

Finally, the analysis accounts for the realization of the initial stop of the past-tense morpheme as follows. It states that this stop is underlyingly unspecified for [voice]. The stop is linked to the [voice]-specification of the preceding segment if this latter segment cannot be specified for [voice] in the output of (lexical) phonology, and otherwise it is specified for the value of [voice] which is the default one in the relevant context.

The analysis incorporates several hypotheses. In order to evaluate the analysis, we tested four of them on the basis of our corpus (Chapters 8 to 10). The first two hypotheses concern the influences of phonological feature specifications and syllabic position at the lexical level on the realization of obstruents, while the third one focusses on the influence of the lexicon. The last hypothesis is especially concerned with obstruent clusters ending in word-initial /d/s. The precise formulation of the four hypotheses can be found below.

> *Hypothesis I*
> Obstruents which are lexically in coda positions are realized as voiced or voiceless independently of their underlying [voice]-specifications, and the phonological feature specifications of the adjacent segments, provided that the realization of these latter specifications does not interfere with the perception of voicing.
>
> *Hypothesis II*
> Obstruents which are lexically in coda position are more likely to be realized as voiced than onset obstruents which are phonologically voiceless, and less likely to be realized as voiced than onset obstruents which are phonologically voiced.
>
> Hypothesis *III*
> There is a systematic difference in realization between word-final obstruents before vowel-initial enclitics in word-combinations which are likely to be retrieved as single units from the lexicon and in word-combinations which are usually computed from their parts.
>
> *Hypothesis IV*
> The realization of a word-final obstruent before a /d/-initial word is determined by the tendency of the /d/-initial word to be hypo-articulated, viz. by its frequency of occurrence, and its contribution to the propositional content of the utterance. If the /d/-initial word is prone to hypo-articulation, its /d/ tends to be realized as voiceless after obstruents, and the preceding obstruent is consequently voiceless as well. Otherwise, the initial /d/ is voiced, and the preceding obstruent is also voiced in the majority of cases.

The hypotheses were tested on the basis of intervocalic single and geminate stops which had been unanimously classified by ear as either voiced or voiceless by two trained phoneticians and the author. Chapter 8 explained why these obstruents were chosen, and extensively discussed the classification method.

It appeared that 85% of the stops were unanimously classified as either voiced or voiceless, and that the probability is less than 0.1% that a stop which was unanimously classified as voiced will be classified as voiceless a second time around, or vice versa. The fact that the sound fragments which were presented to the judges were quite short did not turn out to affect the results of the classification. Finally, an analysis of the closure and burst durations of the stops indicates that the classifications are strongly related to these durations. Nevertheless, a classification method based on the closure and burst durations of the stops cannot yet replace the auditory classification method, since the relation between the classifications and the durations is different for stops in different contexts, and the relation for each context cannot be determined on the basis of the relatively small amount of data discussed in this study (§8.3.5.2).

Chapter 9 discussed the evaluation of *Hypotheses I* to *III* on the basis of 1013 simple intervocalic stops. The data are in accordance with *Hypothesis I* since they show that there is no strong relation between the classifications of word-final stops as voiced or voiceless and the underlying [voice]-specifications of the stops, or the roundness or laxness of the preceding vowels. *Hypothesis II* is confirmed by the data, since they show that intervocalic stops are more likely to be classified as voiced when they are word-medial and underlyingly voiced than when they are word-final, and least when they are word-medial and underlyingly voiceless. Finally, the data show that the word-final stops of the word-combinations *heb ik* 'have I' and *dat ik* 'that I' are more often realized in accordance with their underlying [voice]-specifications. Since *heb ik* and *dat ik* are highly frequent, they are probably more often retrieved as units from the lexicon, and these data suggest that *Hypothesis III* is also correct.

Chapter 10, finally, discussed the testing of *Hypotheses I* and *IV* on the basis of 843 intervocalic stop clusters consisting of a word-final alveolar stop and a word-initial /d/. The classifications of these clusters as voiced or voiceless do not falsify *Hypothesis I*, since they do not show that the realization of these clusters is influenced by the underlying [voice]-specifications of the word-final obstruents. They are in accordance with *Hypothesis IV* in that they suggest that the realization of a cluster is related to the frequency of occurrence of the /d/-initial word, and to the contribution of this word to the propositional content of the utterance.

The four hypotheses could not be tested conclusively because the data set is relatively small, and too little is known about the storage of items in the lexicon and on the contribution of function words to the propositional content of the utterance. Since the hypotheses are in conformity with the data, this study nevertheless makes a reasonable case for the analysis proposed in Chapter 7.

References

Aitchison, J. (1987). *Words in the mind: An introduction to the mental lexicon.* Oxford: Basil Blackwell.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology* **5**, 183-207.

Baumann, M. (1995). *The production of syllables in connected speech.* PhD. dissertation, Catholic University Nijmegen.

Beckman, J.N. (1997). Positional faithfulness, positional neutralisation and Shone vowel harmony. *Phonology* **14**, 1-46.

Beckman, M.E. & J. Edwards (1990). Lengthening and shortening and the nature of prosodic constituency. *Between the grammar and physics of speech* [*Papers in Laboratory Phonology* **1**], edited by J. Kingston & M.E. Beckman, 152-178. Cambridge: Cambridge University Press.

Benua, L. (1995). Identity effects in morphological truncation. *Papers in Optimality Theory* [*University of Massachusetts Occasional Papers in Linguistics* **18**], edited by J. Beckman, L. Walsh Dickey & S. Urbanczyk, 77-136.

Berendsen, E. (1983). Final devoicing, assimilation, and subject clitics in Dutch. *Linguistics in the Netherlands 1993*, edited by H. Bennis & W.U.S. van Lessen Kloeke, 21-30. Dordrecht: Foris.

Berendsen, E. (1986). *The phonology of cliticization.* Dordrecht: ICG Printing.

Berg, R.J.H. van den (1988). *The perception of voicing in Dutch two-obstruent sequences.* PhD. dissertation, Catholic University Nijmegen.

Bergem, D.R. van (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication* **12**, 1-23.

Bergem, D.R. van (1995). *Acoustical and lexical vowel reduction.* Amsterdam: IFOTT.

Blanquaert, E. (1924). Eenheid en nuance in beschaafd-Nederlandse uitspraak. *De Nieuwe Taalgids* **18**, 5-86.

Blancquaert, E. (1934). *Practische uitspraakleer van de Nederlandsche taal*. Antwerpen: De Sikkel.

Boersma, P. (1996). *Praat: Doing phonetics by computer*. Manuscript, University of Amsterdam.

Boersma, P. (1998). *Functional phonology*. The Hague: Holland Academic Graphics.

Booij, G.E. (1981). *Generatieve fonologie van het Nederlands*. Utrecht: Het Spectrum.

Booij, G.E. (1984). French C/∅-alternations, extrasyllabicity and lexical phonology. *The Linguistic Review* **3**, 181-207.

Booij, G.E. (1985). Lexical phonology, final devoicing and subject pronouns in Dutch. *Linguistics in the Netherlands 1985*, edited by H. Bennis & F. Beukema, 21-26. Dordrecht: Foris.

Booij, G.E. (1987). Clitisatie in het Nederlands. *Spektator* **16**, 219-228.

Booij, G.E. (1988). Review of Nespor & Vogel (1986). *Journal of Linguistics* **24**, 515-525.

Booij, G.E. (1995). *The phonology of Dutch*. Oxford: Oxford University Press.

Booij, G.E. (1996). Cliticization as prosodic integration: The case of Dutch. *The Linguistic Review* **13**, 219-242.

Booij, G.E. (1997). Non-derivational phonology meets lexical phonology *Derivations and constraints in phonology*, edited by I. Roca, 261-288. Oxford: Clarendon Press.

Booij, G.E. (1999). The role of the prosodic word in phonotactic generalizations. *Studies on the phonological word*, edited by T.A. Hall & U. Kleinhenz, 47-72. Amsterdam: Benjamins.

Borowsky, T. (forthcoming). Word-faithfulness and the direction of assimilations. To appear in *The Linguistic Review* **17.1**.

Breiman, L., J.H. Friedman, R.A. Olshen & C.J. Stone (1984). *Classification and regression trees*. New York: Chapman & Hall.

Brink, D.T. (1970). *Problems in phonological theory: A generative phonology of Dutch* (University Microfilms International). PhD. dissertation, University of Wisconsin.

Brouwer, D. (1990). Feiten en verzinsels over vrouwen- en mannentaal. *Onze Taal* **58**, 39-42.

Browman, C.P. & L. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* **6**, 201-251.

Browman, C.P. & L. Goldstein (1990a). Tiers in articulatory phonology, with some implications for casual speech. *Between the grammar and physics of speech* [*Papers in Laboratory Phonology* **1**], edited by J. Kingston & M.E. Beckman, 341-376. Cambridge: Cambridge University Press.

Browman, C.P. & L. Goldstein (1990b). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* **18**, 299-320.

Browman, C.P. & L. Goldstein (1992a). Articulatory Phonology: An overview. *Phonetica* **49**, 155-180.

Browman, C.P. & L. Goldstein (1992b). "Targetless" schwa: An articulatory analysis. *Gesture, segment, prosody* [*Papers in Laboratory Phonology* **2**], edited by G.J. Docherty & D.R. Ladd, 26-56. Cambridge: Cambridge University Press.

Browman, C.P. & L. Goldstein (1995). Dynamics and articulatory phonology. *Mind as motion*, edited by T. van Gelder & R.F. Port, 175-193. Cambridge, Massachusetts: MIT Press.

Brown, G. (1977). *Listening to spoken English*. London: Longman.

Bybee, J.L. (1995). *The phonology of the lexicon: Evidence from lexical diffusion*. Paper presented at the Symposium of Usage-Based Models of Language, Rice University, March 1995.

Bybee, J.L. (1996). *Lexicalization of sound change and alternating environments*. Paper presented at the Fifth Conference on Laboratory Phonology, Northwestern University, July 6-8 1996.

Bybee, J.L. & J. Scheibman (1999). The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* **37**, 575-596.

Cambier-Langeveld, T. (1997). The domain of final lengthening in the production of Dutch. *Linguistics in the Netherlands 1997*, edited by J. Coerts & H. de Hoop, 13-24. Amsterdam: Benjamins.

Cambier-Langeveld, T. (1997). *Temporal marking of accents and boundaries*. The Hague: Holland Academic Graphics.

Cho, Y.Y. (1990). A typology of voicing assimilation. *Proceedings of the Ninth West Coast Conference on Formal Linguistics*, edited by A.L. Halpern, 141-155. Leland Stanford Junior University: Centre for the study of language and information.

Chomsky, N. (1964). *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, N. & Halle M. (1968). *The sound pattern of English*. New York: Harper and Row.

Clark, J. & C. Yallop (1995). *An introduction to phonetics and phonology*, second edition. Oxford: Blackwell.

Clements, G.N. & S.J. Keyser (1983). *CV Phonology: A generative theory of the syllable*. Cambridge, Massachusetts : The MIT Press.

Cohn, A.C. (1993). Nasalisation in English: Phonology or phonetics. *Phonology* **10**, 43-81.

Collins, B. & I. Mees (1981). *The sounds of English and Dutch*. Leiden: Leiden University Press.

Cooper, F.S., P.C. Delattre, A.M. Liberman & J.M. Borst (1952). Some experiments on the perception of synthetic sounds. *The Journal of the Acoustical Society of America* **24**, 597-606.

Cucchiarini, C. (1993). *Phonetic transcription: A methodological and empirical study*. PhD. dissertation, Catholic University Nijmegen.

Dalby, J.M. (1984). *Phonetic structure of fast speech in American English* (University Microfilms International). PhD. dissertation, Indiana University.

Debrock, M. (1977). An acoustic correlate of the force of articulation. *Journal of Phonetics* **5**, 61-80.

Delattre, P.C. (1962). Le jeu des transitions de formants et la perception des consonnes. *Proceedings of the Fourth International Congress of Phonetic Sciences held at the University of Helsinky*, edited by A. Sovijaervi & P. Aalto, 407-417. The Hague: Mouton.

Demeulemeester, F. (1962). Assimilatie van stemloze explosieven en frikatieven voor b en d. *Taal en Tongval* **14**, 20-36.

Elias, M. (1977). *Plat-Haags* [*Publicatie van het Instituut voor Algemene Taalwetenschap* **15**]. Amsterdam: University of Amsterdam.

Ernestus, M. (forthcoming). The role of phonology and phonetics in Dutch voice assimilation. *Issues in the phonetics - phonology interface*, edited by V.J. van Heuven, H.G. van der Hulst & J.M. van de Weijer. Under review at publisher.

Fant, G. (1960). *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations*. The Hague: Mouton.

Flemming, E. (1997). Phonetic optimization: Compromise in speech production. *University of Maryland Working Papers in Linguistics* **5**, 72-91.

Forster, K.I. & S.M. Chambers (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behaviour* **12**, 627-635.

Ganong, W.F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* **6**, 110-125.

Garnham, A. (1985). *Psycholinguistics: Central topics*. London: Routledge.

Geerts, G. & H. Heestermans (1992). *Van Dale groot woordenboek der Nederlandse taal*, twelfth edition. Utrecht: Van Dale Lexicografie.

Ginneken, J. van (1935). *Ras en taal*. Amsterdam: N.V. Noord-Hollandsche Uitgevers-Maatschappij.

Goeman, A.C.M. (1999). *T-deletie in Nederlandse dialecten: Kwantitatieve analyse van structurele, ruimtelijke en temporele variatie*. The Hague: Holland Academic Graphics.

Goldsmith, J.A. (1990). *Autosegmental and metrical phonology*. Oxford: Blackwell.

Goldstein, M. & W.R. Dillon (1978). *Discrete discriminant analysis*. New York: John Wiley & Sons.

Grijzenhout, J. (1999). *Duration and voicing in obstruents: Reflections on Dutch and English voice assimilation*. Hand-out Workshop on Laryngeals, Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung, Berlin.

Grijzenhout, J. & M. Krämer (1999). *Final devoicing and voicing assimilation in Dutch derivation and cliticization*. Manuscript, Heinrich-Heine-Universität, Düsseldorf.

Gussenhoven, C. (1986). Over de fonologie van Nederlandse clitica. *Spektator* **15**, 180-200.

Gussenhoven, C. (1988). Intonational phrasing and the prosodic hierarchy. *Phonologica 1988: Proceedings of the 6th International Phonology Meeting*, edited by W.U. Dressler, H.C. Luschützky, O.E. Pfeiffer & J.R. Rennison, 89-99. Cambridge: Cambridge University Press.

Gussenhoven, C. (1989). *Cliticization in Dutch as phonological-word formation*. Manuscript, Catholic University Nijmegen.

Gussenhoven, C. (1992). Dutch. *Journal of the International Phonetic Association* **22**, 45-47.

Gussenhoven, C. (1993). The Dutch foot and the chanted call. *Journal of Linguistics* **29**, 37-63.

Gussenhoven, C. (1999). *Vowel duration, syllable quantity and stress in Dutch*. Manuscript, Catholic University Nijmegen.

Gussenhoven, C. & R.H. Bremmer Jr. (1983). Voiced fricatives in Dutch: sources and present-day usage. *North-Western European Language Evolution* **2**, 55-71.

Gussenhoven, C. & A. Broeders (1976). *The pronunciation of English: A course for Dutch learners*. Groningen: Wolters-Noordhoff-Longman.

Guy, G.R. (1980). Variation in the group and the individual: The case of final stop deletion. *Locating language in time and space*, edited by W. Labov, 1-36. New York: Academic Press.

Haeringen, C.B. van (1955). Is tat juist, op tie manier. *Taal en Tongval* **7**, 49-57.

Haeringen, C.B. van (1971). Wat vin je hiervan? *Taal en Tongval* **12**, 161-168.

Hale, M. & C. Reiss (forthcoming). Substance abuse and dysfunctionalism: Current trends in phonology. *Linguistic Inquiry*.

Halle, M., G.W. Hughes & J.P.A. Radley (1957). Acoustic properties of stop consonants. *The Journal of the Acoustical Society of America* **29**, 107-116.

Hayes, B. (1986). Inalterability in CV Phonology. *Language* **62**, 321-351.

Hinkens, F. (1992). *Dialect levelling, structural and sociolinguistic aspects*. PhD. dissertation, Catholic University Nijmegen.

Kirsner, R.S. & V.J. van Heuven (1988). The significance of demonstrative position in Modern Dutch. *Lingua* **76**, 265-304.

Kirsner, R.S., V.J. van Heuven & J.F.M. Vermeulen (1987). Text-type, context and demonstrative choice in written Dutch: some experimental data. *Text* **7**, edited by A. van Dijk, 117-144. Amsterdam: Mouton de Gruyter.

Hulst, H. van der (1984). *Syllable structure and stress in Dutch*. Dordrecht: Foris.

Hulst, H. van der (1985). Ambisyllabicity in Dutch. *Linguistics in the Netherlands 1985*, edited by H. Bennis & F. Beukema, 57-66. Dordrecht: Foris.

International Phonetic Association (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Itô, J. (1989). A prosodic theory of epenthesis. *Natural Language and Linguistic Theory* 7, 217-259.

Jongenburger, W. & V. van Heuven (1991). The distribution of (word initial) glottal stop in Dutch. *Linguistics in the Netherlands 1991*, edited by F. Drijkoningen & A. van Kemenade, 101-110. Amsterdam: Benjamins.

Jongenburger, W. & V. van Heuven (1993). Sandhi processes in natural and synthetic speech. *Analysis and synthesis of speech: strategic research towards high-quality text-to-speech generation*, edited by V.J. van Heuven & L.C.W. Pols, 261-276. Berlin: Mouton de Gruyter.

Jongman, A., J.A. Sereno, M. Raaijmakers & A. Lahiri (1992). The phonological representation of [voice] in speech perception. *Language and Speech* **35**, 137-152.

Jun, A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. *Phonology and phonetic evidence* [*Papers in Laboratory Phonology* **4**], edited by B. Connel & A. Arvaniti, 235-253. Cambridge: Cambridge University Press.

Kager, R. (1989). *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.

Kaiser, L. (1958). Onderzoek naar assimilatieverschijnselen. *Album Edgard Blancquaert: De gehuldigde aangeboden ter gelegenheid van zijn emeritaat door kollega's vakgenoten en oud-leerlingen*, 31-41. Tongeren: George Michels.

Kaisse, E.M. (1985). *Connected speech: The interaction of syntax and phonology*. Orlando, Florida: Academic Press.

Keating, P. (1985). Universal Phonetics and the organization of grammars. *Phonetic linguistics: Essays in honor of Peter Ladefoged*, edited by V. Fromkin, 115-132. Orlando: Academic Press.

Keating, P. (1988a). The phonology-phonetics interface. *The Cambridge survey*. Vol. I*: Linguistic theory*, edited by F.J. Newmeyer, 281-302. Cambridge: Cambridge University Press.

Keating, P. (1988b). Underspecification in phonetics. *Phonology* **5**, 275-292.

Keating, P. (1990a). Phonetic representations in a generative grammar. *Journal of Phonetics* **18**, 321-334.

Keating, P. (1990b). The window model of coarticulation: Articulatory evidence. *Between the grammar and physics of speech* [*Papers in Laboratory Phonology* **1**], edited by J. Kingston & M.E. Beckman, 451-470. Cambridge: Cambridge University Press.

Keating, P. (1998). Word-level phonetic variation in large speech corpora. *Papers of the conference on "the word as a phonetic unit"* [*ZAS papers in linguistics* **11**], edited by A. Alexiadou, N. Fuhrhop, U. Kleinhenz & P. Law, 35-50. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung.

Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman.

Kenstowicz, M. (1996). Base-identity and uniform exponence: Alternatives to cyclicity. *Current trends in phonology: Models and methods*, edited by J. Durand & B. Laks, 363-393. Salford, Manchester: Salford University Press, European Studies Research Institute.

Kenstowicz, M. & C. Kisseberth (1979). *Generative phonology: Description and theory*. New York: Academic Press.

Kingston, J. & R.L. Diehl (1994). Phonetic knowledge. *Language* **70**, 419-454.

Kiparsky, P. (1982). From cyclic phonology to lexical phonology. *The structure of phonological representations* Volume I, edited by H. van der Hulst & N. Smith, 131-175. Dordrecht: Foris.

Kiparsky, P. (1997). *The phonology-morphology interface*. Course 491.21, LSA Summer Institute, Ithaca.

Kirsner, R.S. & V.J. van Heuven (1988). The significance of demonstrative position in Modern Dutch. *Lingua* **76**, 265-304.

Kirsner, R.S., V.J. van Heuven & J.F.M. Vermeulen (1987). Text-type, context and demonstrative choice in written Dutch: Some experimental data. *Text* **7**, 117-144.

Koefoed, G.A.T. (1979). Paradigmatische invloeden op fonetische processen. *Morfologie in Nederland* [*Glot Special*], edited by T. Hoekstra & H. van der Hulst, 53-70. Leiden: Leiden University.

Kohler, K.J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. *Speech production and speech modelling*, edited by W.J. Hardcastle & A. Marchal, 21-33. Dordrecht: Kluwer Academic Publishers.

Kohler, K.J. (1998). The disappearance of words in connected speech. *Papers of the conference on "the word as a phonetic unit"* [*ZAS papers in linguistics* **11**], edited by A. Alexiadou, N. Fuhrhop, U. Kleinhenz, and P. Law, 21-33. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1975). *What is a linguistic fact?* Lisse: The Peter de Ridder Press.

Labov, W. (1994). *Principles of linguistic change*. Vol. 1: *Internal factors*. Oxford: Blackwell.

Ladefoged, P. & I. Maddieson (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Leenen, J. (1954). Een Limburgse en Nederlandse uitspraakregel. *Taal en Tongval* **6**, 1-24.

Leenen, J. (1955). De oorsprong van de Limburgse en Nederlandse uitspraakregel. *Taal en Tongval* **7**, 58-66.

Lehiste, I. & G. Peterson (1959). Vowel amplitude and phonemic stress in American English. *The Journal of the Acoustical Society of America* **31**, 428-435.

Levelt, W.J.M. (1972). Some psychological aspects of linguistic data. *Linguistische Berichte* **17**, 18-30.

Levelt, W.J.M. (1989). *Speaking: From intention to articulation*. Cambridge, Massachusetts: MIT Press.

Liebetrau, A.M. (1983). *Measures of association*. Beverly Hills, California: Sage.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. *Speech production and speech modelling*, edited by W.J. Hardcastle & A. Marchal, 403-439. Dordrecht: Kluwer Academic Publishers.

Lisker, L. (1957). Closure duration and the intervocalic voiced-voiceless distinction in English. *Language* **33**, 42-49.

Lisker, L. & A.S. Abramson (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* **20**, 384-422.

Lombardi, L. (1995a). Laryngeal features and privativity. *The Linguistic Review* **12**, 35-49.

Lombardi, L. (1995b). Restrictions on direction of voicing assimilation: An OT account. *University of Maryland Working Papers in Linguistics* **3**, 89-115.

Lombardi, L. (1995c). *Why place and voice are different: Constraint interactions and featural faithfulness in Optimality Theory*. Manuscript, University of Maryland.

Lombardi, L. (1996). Postlexical rules and the status of privative features. *Phonology* **13**, 1-38.

Lombardi, L. (1999). Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language & Linguistic Theory* **17**, 267-302.

Losiewicz, B.L. (1992). *The effect of frequency on linguistic morphology*. Ph.D dissertation, University of Texas.

Manuel, S.Y. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America* **88**, 1286-1298.

Marsi, E.C., P.A.J.M. Coppen, C.H.M. Gussenhoven & A.C.M. Rietveld (1997). Prosodic and intonational domains in speech synthesis. *Progress in speech synthesis*, edited by J.P.H. van Santen, R.W. Sproat, J.P. Olive & J. Hirschberg, 477-493. New York: Springer.

Martens, L. & H. Quené (1994). Degemination of Dutch fricatives in three different speech rates. *Linguistics in the Netherlands 1994*, edited by R. Bok-Bennema & C. Cremers, 119-126. Amsterdam: Benjamins.

Martin, W. (1968). De verdoffing van gedekte en ongedekte e in niet-hoofdtonige positie bij Romaanse leenwoorden in het Nederlands. *De Nieuwe Taalgids* **61**, 162-181.

Mascaró, J. & L. Wetzels (1999). *The typology of voicing and devoicing*. Manuscript, Vrije Universiteit Amsterdam & Universitat Autonoma de Barcelona.

McCarthy, J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry* **12**, 373-418.

McCarthy, J. (1998). *Sympathy & phonological opacity*. Manuscript, University of Massachusetts.

McCarthy, J. & A. Prince (1993a). *Prosodic morphology I: Constraint interaction and satisfaction*. Manuscript, University of Massachusetts & Rutgers University.

McCarthy, J. & A. Prince (1993b). Generalized alignment. *Yearbook of Morphology 1993*, edited by G.E. Booij & J. van Marle, 79-153. Dordrecht: Kluwer.

McCarthy, J. & A. Prince (1995). Faithfulness and reduplicative identity. *Papers in Optimality Theory* [*University of Massachusetts Occasional Papers in Linguistics* **18**], edited by J.N. Beckman, L. Walsh Dickey, & S. Urbanczyk, 249–384.

Meinsma, G.L. (1958). Assimilatie en assimilatie-onderzoek. *Taal en Tongval* **10**, 107-114.

Menert, L. (1994). *Experiments on voice assimilation in Dutch: Prosodic structures and tempo*. PhD. dissertation, University of Utrecht.

Mester, R.A. & J. Itô (1989). Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* **65**, 258-293.

Miller, J.L. (1981). Some effects of speaking rate on phonetic perception. *Phonetica* **38**, 159-180.

Model, J. (1991). *Grammatische analyse*. Dordrecht: Foris.

Mohanan, K.P. (1986). *The theory of lexical phonology*. Dordrecht: D. Reidel Publishing Company.

Nerbonne, J. (1999). Edit distance and dialect proximity. *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison*, edited by D. Sankoff & J. Kruskal, v-xv. Leland Stanford Junior University: Centre for study of Language and Information.

Nespor, M. (1990). Vowel deletion in Italian: The organization of the phonological component. *The Linguistic Review* **7**, 375-398.

Nespor, M. & I. Vogel (1982). Prosodic domains of external sandhi rules. *The structure of phonological representations* part I, edited by H. van der Hulst & N. Smith, 225-255. Dordrecht: Foris Publications.

Nespor, M. & I. Vogel (1986). *Prosodic phonology*. Dordrecht: Foris.

Neu, H. (1980). Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. *Locating language in time and space*, edited by W. Labov, 37-54. New York: Academic Press.

Newmeyer, F.J. (1983). *Grammatical theory: Its limits and its possibilities*. Chicago: The University of Chicago Press.

Nooteboom, S.G. (1972). *Production and perception of vowel duration: A study of the durational properties of vowels in Dutch*. PhD. dissertation, University of Utrecht.

Nooteboom, S.G. (1979). Perceptual adjustment to speech rate: A case of backward perceptual normalization. *Anniversaries in phonetics: Studia gratulatoria dedicated to Hendrik Mol*, edited by H.V. Deighton-van Witsen, 225-269. Amsterdam: Institute of Phonetic Sciences.

Nooteboom, S.G. & G.J.N. Doodeman (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America* **67**, 276-287.

Nooteboom, S.G. & I.H. Slis (1969). A note on rate of speech. *IPO Annual Progress Report*, 58-60. Eindhoven: Institute for Perception Research.

Ohala, J.J. (1981). Speech timing as a tool in phonology. *Phonetica* **38**, 204-212.

Ohala, J.J. (1983). The origin of sound patterns in vocal tract constraints. *The production of speech*, edited by P.F. MacNeilage, 189-216. Berlin: Springer.

Ohala, J.J. (1986). Consumer's guide to evidence in phonology. *Phonology Yearbook* **3**, 3-26.

Ohala, J.J. (1990a). The phonetics and phonology of aspects of assimilation. *Between the grammar and physics of speech* [*Papers in Laboratory Phonology* **1**], edited by J. Kingston & M.E. Beckman, 258-275. Cambridge: Cambridge University Press.

Ohala, J.J. (1990b). There is no interface between phonology and phonetics: A personal view. *Journal of Phonetics* **18**, 153-171.

Ohde, R.N. (1984). Fundamental frequency as an acoustic correlate of stop consonant voicing. *Journal of the Acoustical Society of America* **75**, 224-230.

Oostendorp, M. van (1995). *Vowel quality and syllable projection*. PhD. dissertation, Catholic University Brabant.

Oostendorp, M. van (1997). Style levels in conflict resolutions. *Variation, change and phonological theory*, edited by F. Hinskens, R. van Hout & W. L. Wetzels, 207-230. Amsterdam: Benjamins.

Oss, F. van & C. Gussenhoven (1984). De Nederlandse slot-n in het nieuws. *Gramma* **8**, 37-45.

Ottow-Kolman, A. (1989). T-deletie na sonoranten in Nederlandse dialecten. *Taal en Tongval* **41**, 67-85.

Pater, J. (1997). Minimal violation and phonological development. *Language Acquisition* **6**, 201-253.

Peperkamp, S. (1997). *Prosodic words*. The Hague: Holland Academic Graphics.

Pierrehumbert, J. (1990). Phonological and phonetic representation. *Journal of Phonetics* **18**, 375-394.

Prince, A. & P. Smolensky (1993). *Optimality Theory: Constraint interaction in Generative Grammar*. Manuscript, Rutgers University & University of Colorado at Boulder.

Reenen, P. Th. van (1986). The vocalization of /l/ in standard Dutch, a pilot study of an ongoing change. *Linguistics in the Netherlands 1986*, edited by F. Beukeman & A. Hulk, 189-197. Dordrecht: Foris.

Reenen, P. Th. van (1994). Driemaal /r/ in de Nederlandse dialecten. *Taal en Tongval* **46**, 54-72.

Reenen, P. Th. van & A. Coetzee (1996). Afrikaans, a daughter of Dutch. *The origins and development of emigrant languages: Proceedings from the second Rasmus Rask Colloquium, Odense University, Novermber 1994*, edited by H.F. Nielsen & L. Schøsler, 71-101. Odense: Odense University Press.

Reenen, P. Th. van & M. Elias (1998). *Taalverschillen: Een werkboek over variatie en verandering in taal*. Bussum: Coutinho.

Reenen, P. Th. van & E. Wattel (1992). De uitspraak van /s/ en /z/ voor klinker in het Nederlands: Zes eeuwen variatie. *De binnenbouw van het Nederlands: Een bundel artikelen voor Piet Paardekoper*, edited by H. Bennis & J.W. de Vries (eds.), 291-306. Dordrecht: ICG Publications.

Rietveld, A.C.M. & V.J. van Heuven (1997). *Algemene fonetiek*. Bussum: Coutinho.

Rietveld, T. & R. van Hout (1993). *Statistical techniques for the study of language and language behaviour*. Berlin: Mouton de Gruyter.

Rietveld, T., J. Kerkhoff & C. Gussenhoven (1999). Prosodic structure and vowel duration in Dutch. *Proceedings of the 14th International Congress of Phonetic Sciences, San Francisco, 1-7 August 1999*, edited by J.J. Ohala, Y. Hasegawa, M. Ohala, D. Granville & A.C. Baily, 463-466. Berkeley: Linguistic Department, University of California.

Rietveld, A.C.M. & F.J. Koopmans-van Beinum (1987). Vowel reduction and stress. *Speech Communication* **6**, 217-229.

Rijnbach, A. van & W.F. Kramer (1939). De assimilatie van stem en plaats in het Nederlandsch bij het lezen van losse zinnen. *Onze Taaltuin* **8**, 97-107.

Rodgers, J. (1998). *Segmental and suprasegmental influences on the realization of voicing in English vowels*. Paper presented at the Sixth Conference on Laboratory Phonology York, 2-4 July.

Ross, S. (1976). *A first course in probability*. New York: Macmillan Publishing Company.

Rubenstein, H., L. Garfield & J. Millikan (1970). Homographic entries in the internal lexicon. *Journal of Verbal Learning and Verbal Behavior* **9**, 487-499.

Saussure, F. de (1916). *Cours de linguistique générale*. Page reference to the second edition publiced by C. Bally & A. Sechehaye. Paris: Payot & C^{ie}, 1922.

Selkirk, E. (1980). Prosodic domains in phonology: Sanskrit revisited. *Juncture*, edited by M. Aronoff & M.L. Kean, 107-129. Saratoga, Calif.: Anma Libri & Co.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook* **3**, 371-405.

Selkirk, E. (1995). The prosodic structure of function words. *Papers in Optimality Theory* [*University of Massachusetts Occasional Papers in Linguistics* **18**], edited by J.N. Beckman, L. Walsh Dickey & S. Urbanczyk, 439-469. Amherst: GLSA.

Siegel, S. & N.J. Catellan (1988). *Nonparametric statistics for the behavioral sciences*, second edition. New York: McGraw-Hill.

Slis, I.H. (1982). Assimilatie van stem in het Nederlands. *Glot* **5**, 235-261.

Slis, I.H. (1983). Assimilation of voice as a function of stress and word boundaries. *Proceedings of the Institute of Phonetics Nijmegen* **7**, 74-82.

Slis, I.H. (1985). *The voiced-voiceless distinction and assimilation of voice in Dutch*. Helmond: Wibro.

Slis, I.H. (1987). Assimilation of voice in Dutch. *In honour of Ilse Lehiste*, edited by R. Channon & L. Stockey, 225-239. Dordrecht: Foris.

Slis, I.H. & A. Cohen (1969). On the complex regulating the voiced-voiceless distinction II. *Language and Speech* **12**, 80-102; 137-155.

Slis, I.H. & M. van Heugten (1989). Voiced-voiceless distinction in Dutch fricatives. *Linguistics in the Netherlands 1989*, edited by H. Bennis & A. van Kemenade, 123 -132. Dordrecht: Foris.

Slis, I.H. & D. van Nierop (1970). On the forward masking threshold of vowels in VC-combinations. *IPO Annual Progress Report*, 68-73. Eindhoven: Institute for Perception Research.

Solomon, R.L. & D.H. Howes (1951a). Word frequency, personal values, and visual duration thresholds. *Psychological Review* **58**, 256-270.

Solomon, R.L. & D.H. Howes (1951b). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology* **41**, 401-410.

Steinberg, D. & P. Colla (1984). *CART: Tree-structured nonparametric data analysis*. San Diego, California: Salford Systems.

Stemberger, J.P. & B. MacWhinney (1988). Are inflected forms stored in the lexicon? *Theoretical morphology: Approaches in modern linguistics*, edited by M. Hammond & M. Noonan, 101-116. San Diego: Academic Press.

Steriade, D. (1987). Redundant values. *Papers from the 23d Annual Regional Meeting of the Chicago Linguistic Society. Part two: Parasession on Autosegmental and Metrical Phonology*, 339-362. Chicago: Chicago Linguistic Society.

Stroop, J.P.A. (1974). Iets over de uitspraak van protonische vokaal in Romaanse leenwoorden. *De Nieuwe Taalgids* **67**, 314-30.

Stroop, J.P.A. (1998a). Wordt het Poldernederlands model? *Noordzee taal & letteren* **1**, 11-13.

Stroop, J.P.A. (1998b). *Poldernederlands: Waardoor het ABN verdwijnt*. Amsterdam: Bert Bakker.

Trommelen, M.T.G. (1983). *The syllable in Dutch: With special reference to diminutive formation*. Dordrecht: Foris.

Trommelen, M. & W. Zonneveld (1979). *Inleiding in de generatieve fonologie*. Muiderberg: Coutinho.

Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: University Press.

Uit den Boogaart, P.C. (1975). *Woordfrequenties in geschreven en gesproken Nederlands*. Utrecht: Oosthoek, Scheltema, and Holkema.

Van de Velde, H. (1996). *Variatie en verandering in het gesproken standaard-Nederlands (1935-1993)*. PhD. dissertation, Catholic University Nijmegen.

Van de Velde, H. & R. van Hout (1999). The pronunciation of (r) in Standard Dutch. *Linguistics in the Netherlands 1999*, edited by R. van Bezooijen & R. Kager, 177-188. Amsterdam: Benjamins.

Vieregge, W.H. (1987). Basic aspects of phonetic segmental transcription. *Probleme der Phonetischen Transkription*, edited by A. Almeida & A. Braun, 5-55. Stuttgart: Franz Steiner Verslag Wiesbaden GMBH.

Vieregge, W.H. & A.P.A. Broeders (1993). Intra- and interspeaker variation of /r/ in Dutch. *Proceedings of Eurospeech 1993*, Volume 1, 267-270. Edinburgh: CEP Consultants.

Vogel, I. (1994). Phonological interfaces in Italian. *Issues and theory in Romance linguistics: Selected papers from the linguistic symposium on Romance languages XXIII*, edited by M.L. Mazzola, 109-126. Washington: Georgetown University Press.

Voortman, B. (1994). *Regionale variatie in het taalgebruik van notabelen: Een sociolinguïstisch onderzoek in Middelburg, Roermond en Zutphen*. PhD. dissertation, University of Amsterdam.

Vries, J.W. de (1974). De slot-t in consonantclusters te Leiden: Een sociolinguïstisch onderzoek. *Forum der Letteren* **15**, 235-250.

Waals, J. (1999). *An experimental view of the Dutch syllable*. The Hague: Holland Academic Graphics.

Warren, R.M. (1976). Auditory illusions and perceptual processes. *Contemporary issues in experimental phonetics*, edited by N.J. Lass, 389-417. New York: Academic Press.

Westbury, J.R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* **73**, 1322-1336.

Wetzels, L. (1982). *Analogie et lexicque: le problème de l'opacité en phonologie générative*. PhD. dissertation, Catholic University Nijmegen.

Wetzels, L. (1994). *De fonologie van stem.* Manuscript Vrije Universiteit Amsterdam.

Wetzels, L. (1997). Bakairi and the feature "voice". *Abralin* **20**, 21-34.

Wightman, C.W., S. Shattuck-Hufnagel, M. Ostendorf & P.J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* **91**, 1707-1717.

Wijk, C. van & G. Kempen (1980). Funktiewoorden: Een inventarisatie voor het Nederlands. *ITL: Review of Applied Linguistics* **47**, 53-68.

Wijst, P. van der (1996). *Politeness in requests and negations*. PhD. dissertation, Catholic University Brabant.

Zonneveld, W. (1978). *A formal theory of exceptions in generative phonology*. Lisse: Peter de Ridder Press.

Zonneveld, W. (1982). The descriptive power of the Dutch theme-vowel. *Spektator* **11**, 342-365.

Zonneveld, W. (1983). Lexical and phonological properties of Dutch devoicing assimilation. *Sound structures: Studies for Antonie Cohen*, edited by M. van den Broecke, V. van Heuven & W. Zonneveld, 297-312. Dordrecht: Foris.

Zsiga, E.C. (1992). A mismatch between morphological and prosodic domains: Evidence from two Igbo rules. *Phonology* **9**, 101-135.

Zsiga, E.C. (1995). An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. *Phonology and phonetic evidence* [*Papers in Laboratory Phonology* **4**], edited by B. Cornell & A. Arvanti, 282-302. Cambridge: Cambridge University Press.

Zsiga, E.C. (1997). Features, gestures, and Igbo vowels: An approach to the phonology-phonetics interface. *Language* **73**, 227-274.

Zwaardemaker, H. & L.P.H. Eijkman (1928). *Leerboek der phonetiek*. Haarlem: De Erven F. Bohn.

Zwicky, A.M. (1972). On casual speech. *Papers from the Eighth Regional Meeting Chicago Linguistic Society*, edited by P.M. Peranteau, J.N. Levi & G.C. Phares, 607-615. Chicago: Chicago Linguistic Society.

Appendices

A Information on Subjects A to P

The infomation was valid in 1995, 1996.

- Subject A

  Year of birth: 1966.

  Former and present places of residence: Hoorn, Houten, Haarlem, Utrecht.

  Profession: PhD. student of applied linguistics syntax/morphology.

- Subject B

  Year of birth: 1947.

  Former and present places of residence: Amsterdam, Weesp, Ouderkerk aan de Amstel.

  Profession: professor of applied linguistics syntax/morphology.

- Subject C

  Year of birth: 1963.

  Former and present places of residence: Hoorn, Amsterdam.

  Profession: PhD. student of general literature.

- Subject D

  Year of birth: 1965.

  Former and present places of residence: Haarlem, Amsterdam.

  Profession: PhD. student of general literature.

- Subject E

  Year of birth: 1967.

  Former and present places of residence: Amstelveen, Amsterdam.

  Profession: Sociologist, and bookkeeper.

- Subject F

  Year of birth: 1941.

  Former and present places of residence: Voorburg, The Hague, Haarlem, Amstelveen.

  Profession: professor of computational linguistics.

- Subject G

  Year of birth: 1948.

  Former and present places of residence: Schiedam, Leiden, Alphen aan de Rijn.

  Profession: professor of the history of Dutch linguistics.

- Subject H

  Year of birth: 1971.

  Place of birth and place of residence: Amsterdam.

  Profession: doctor of medicine.

- Subject I

  Year of birth: 1952.

  Former and present places of residence: Amsterdam, Naarden, Amsterdam.

  Profession: specialized librarian general linguistics and general literature.

- Subject J

  Year of birth: 1957.

  Former and present places of residence: Hillegom, Amsterdam.

  Profession: archaeologist.

- Subject K

  Year of birth: 1959.

  Former and present places of residence: Amsterdam, Osdorp, Amsterdam.

  Profession: post-doctoral researcher in Dutch syntax.

- Subject L

  Year of birth: 1961.

  Former and present places of residence: Haarlem, Amsterdam.

  Profession: PhD. student of historical Dutch syntax.

- Subject M

  Year of birth: 1955.

  Former and present places of residence: Broek op Langedijk, Haarlem, Amsterdam.

  Profession: teacher of computer sciences.

- Subject N

    Year of birth: 1967.

    Former and present places of residence: Haarlem, Santpoort, Amsterdam.

    Profession: PhD. student of historical of English syntax.

- Subject O

    Year of birth: 1962.

    Former and present places of residence: Haarlem, Amsterdam.

    Profession: technical writer.

- Subject P

    Year of birth: 1960.

    Former and present places of residence: Hilversum, Huizen, Amsterdam.

    Profession: Economist.

B Intuitions on intervocalic stops

The intuitions were elicited in May 1997

Table B.1 Subject A's intuitions on the realization of intervocalic coronal stops as [t] or [d], broken down by item type. The horizontal lines demarcate the items which are words and the items which are word-combinations and end in the same function word.

Item type			Possible realization of the stop
letter	/ˈlɛt̠ər/	'letter'	[t]
meter	/ˈmet̠ər/	'meter'	[t]
ontzettend	/ɔntˈzɛt̠ənd/	'terrible'	[t]
water	/ˈʋat̠ər/	'water'	[t] and [d]
moeten	/ˈmut̠-ən/	'must-inf.'	[t]
praten	/ˈprat̠-ən/	'talk-inf'	[t]
zetten	/ˈzɛt̠-ən/	'put-inf.'	[t]
mede	/ˈmed̠ə/	'co-'	[d]
vader	/ˈvad̠ər/	'father'	[d]
laat ik	/ˈlat̠ ɪk/	'let I'	[t] and [d]
moet ik	/ˈmut̠ ɪk/	'must I'	[t] and [d]
weet ik	/ˈʋet̠ ɪk/	'know I'	[d]
zit ik	/ˈzɪt̠ ɪk/	'sit I'	[t] and [d]
bied ik	/ˈbid̠ ɪk/	'offer I'	[d]
had ik	/ˈhɑd̠ ɪk/	'had I'	[d]
mijd ik	/ˈmɛid̠ ɪk/	'avoid I'	[d]
moet het	/ˈmut̠ ət/	'must it'	[t] and [d]
ziet het	/ˈzit̠ ət/	'sees it'	[t] and [d]
bied het	/ˈbid̠ ət/	'offer it'	[d]
deed het	/ˈded̠ ət/	'did it'	[d]
doet er	/ˈdut̠ ər/	'does there'	[t]
ziet er	/ˈzit̠ ər/	'sees there'	[t]
had er	/ˈhɑd̠ ər/	'had there'	[t] and [d]
weet ie	/ˈwet̠ i/	'knows he'	[t]
ziet ie	/ˈzit̠ i/	'sees he'	[t]
had ie	/ˈhɑd̠ i/	'had he'	[t] and [d]

Table B.2 Intuitions of Subjects B to F on the realization of intervocalic coronal stops as [t] or [d], broken down by item type. The horizontal lines demarcate the items which are words and the items which are word-combinations and end in the same function word. The meanings and underlying representations of the items can be found in Table B.1.

Item type	Possible realization of the stop according to Subject				
	B	C	D	E	F
letter	[t]	[t]	[t]	[t]	[t] and [d]
meter	[t] and [d]	[t]	[t]	[t]	[t]
ontzettend	[t] and [d]	[t]	[t]	[t]	[d]
water	[t]	[t]	[t]	[t]	[t]
moeten	[t]	[t]	[t]	[t]	
praten	[t] and [d]	[t]	[t]	[t]	[t]
zetten	[t] and [d]	[t]	[t]	[t]	[t] and [d]
mede	[d]	[d]	[d]	[d]	[d]
vader	[d]	[d]	[d]	[d]	[d]
laat ik	[d]	[t] and [d]	[t]	[t]	[d]
moet ik	[d]	[t] and [d]	[t]	[t] and [d]	[d]
weet ik	[d]	[t] and [d]	[t]	[d]	[d]
zit ik	[d]	[t] and [d]	[t]	[t]	[d]
bied ik	[d]	[t] and [d]	[t]	[t]	[d]
had ik	[d]	[t] and [d]	[t] and [d]	[t]	[d]
mijd ik	[d]	[t] and [d]	[t]	[t]	[d]
moet het	[t] and [d]	[t] and [d]	[t]	[t]	[t] and [d]
ziet het	[t] and [d]	[t] and [d]	[t]	[t]	[d]
bied het	[t] and [d]	[t] and [d]	[t]	[t] and [d]	[t] and [d]
deed het	[t] and [d]	[t] and [d]	[t]	[t]	
doet er	[t] and [d]	[t] and [d]	[t]	[t]	[t] and [d]
ziet er	[t] and [d]	[t] and [d]	[t]	[t]	[t]
had er	[t] and [d]	[t] and [d]	[t]	[t]	[t] and [d]
weet ie	[t] and [d]	[t] and [d]	[t]	[t]	[t]
ziet ie		[t] and [d]	[t]	[t]	[t]
had ie	[t]	[t] and [d]	[t]	[t]	[t] and [d]

Table B.3 Intuitions of Subjects G to K on the realization of intervocalic coronal stops as [t] or [d], broken down by item type. The horizontal lines demarcate the items which are words and the items which are word-combinations and end in the same function word. The meanings and underlying representations of the items can be found in Table B.1.

Item type	Possible realization of the stop according to Subject				
	G	H	I	J	K
letter	[t]	[t]	[t]	[t]	[t]
meter	[t]	[t] and [d]	[t]	[t]	[t]
ontzettend	[t]	[t]	[t] and [d]	[t]	[t] and [d]
water	[t]	[t] and [d]	[t]	[t]	[t]
moeten	[t]	[t]	[t]	[t]	[t]
praten	[t]	[t]	[t] and [d]	[t]	[t]
zetten	[t]	[t]	[t]	[t]	[t]
mede	[d]	[d]	[d]	[d]	[d]
vader	[d]	[d]	[d]	[d]	[d]
laat ik	[t]	[t]	[t] and [d]	[t]	[t]
moet ik	[t]	[d]	[t]	[t]	[t]
weet ik	[t] and [d]	[d]	[t] and [d]	[t]	[t]
zit ik	[t]	[t] and [d]	[t]	[t]	[t]
bied ik	[t] and [d]	[t]	[t] and [d]	[t]	[t] and [d]
had ik	[d]	[d]	[t] and [d]	[t]	[t] and [d]
mijd ik	[d]	[t]	[t]	[t]	[t] and [d]
moet het	[t]	[t] and [d]	[t]	[t]	[t]
ziet het	[t]	[d]	[t]	[t]	[t]
bied het	[t]	[d]	[t]	[t]	[d]
deed het	[t]	[d]	[t]	[t]	[d]
doet er	[t]	[t] and [d]	[t] and [d]	[t]	[t] and [d]
ziet er	[t]	[t] and [d]	[t] and [d]	[t]	[t]
had er	[t]	[t] and [d]	[t] and [d]	[t]	[t] and [d]
weet ie	[t]	[t]	[t]	[t]	[t]
ziet ie	[t]	[t]	[t]	[t]	[t]
had ie	[t]	[t]	[t]	[t]	[t] and [d]

Table B.4 Intuitions of Subjects L to P on the realization of intervocalic coronal stops as [t] or [d], broken down by item type. The horizontal lines demarcate the items which are words and the items which are word-combinations and end in the same function word. The meanings and underlying representations of the items can be found in Table B.1.

Item type	Possible realization of the stop according to Subject				
	L	M	N	O	P
letter	[t]	[t]	[t]	[t]	[t]
meter	[t]	[t]	[t]	[t]	[t]
ontzettend	[t]	[t]	[t]	[t]	[t]
water	[t]	[t]	[t]	[t]	[t]
moeten	[t]	[t]	[t]	[t]	[t]
praten	[t]	[t]	[t]	[t]	[t]
zetten	[t]	[t]	[t]	[t]	[t]
mede	[d]	[d]	[d]	[d]	[t]
vader	[d]	[d]	[d]	[d]	[t]
laat ik	[t]	[t] and [d]	[t]	[t] and [d]	[t]
moet ik	[t]	[t] and [d]	[t]	[t] and [d]	[t]
weet ik	[t]	[t] and [d]	[t]	[t] and [d]	[d]
zit ik	[t]	[t] and [d]	[t]	[t] and [d]	[t]
bied ik	[t] and [d]	[t] and [d]	[t]	[t] and [d]	[t]
had ik	[t] and [d]	[t] and [d]	[t]	[t] and [d]	[d]
mijd ik	[t]	[t]	[t]	[t] and [d]	[d]
moet het	[t]	[t]	[t]	[t]	[t]
ziet het	[t]	[t]	[t]	[t]	[t]
bied het	[t]	[t]	[t]	[t] and [d]	[t]
deed het	[t]	[t]	[t]	[t]	[d]
doet er	[t]	[t]	[t]	[t]	[t] and [d]
ziet er	[t]	[t]	[t]	[t] and [d]	[t]
had er	[t]	[t]	[t]	[t]	[d]
weet ie	[t]	[t]	[t]	[t]	[t]
ziet ie	[t]	[t]	[t]	[t]	[t]
had ie	[t]	[t]	[t]	[t]	[d]

C Instructions for the role-play

The instructions were originally in Dutch

C.1 Instructions for the store owner

One of your friends is about to visit you. The last time you saw him was at a party some days ago. You met many old friends at that party, and heard many fierce discussions (e.g. about euthanasia, the education policy, religion). You and your friend will talk some time about that **party** (the people, the discussions, the food).

Then, you will ask your friend how his driving test went. He told you at the party that he had to take his driving test very soon, and you want to know the result. Moreover, you yourself have some stories about driving lessons and driving tests.

Finally, you will talk about what you have been doing since the party (amusement parks, holidays, visits to family/friends, TV-programmes, pets, cinema?).

After you have talked about several topics for minimally a quarter of an hour, your friend will tell you that he has not come only for pleasure. He wants to sell you some products.

Your friend is the salesman of **Campex**. Campex is a firm producing camping articles. It is famous for the high quality of its products. The quality of its **sleeping-bags**, however, is not very high.

You are the owner of the shop **Sports & Travel**. Sports & Travel sells all kinds of articles in the area of sports and leisure time. Camping articles also belong to the assortment.

You bought **back-packs type A** and **sleeping-bags type A** from your friend last spring. Your stock has to be replenished urgently with these same products. You will therefore negotiate with your friend. As it is getting on for the end of the camping season, you **do not want to buy lots of more than 75 pieces**. Moreover, you only want to buy back-packs and sleeping-bags.

Of course, you will try to pay as little as possible for every product, mostly by bargaining for **reductions**. When fixing the prices, you will have to keep in mind at which prices the products are sold in your shop. **The purchase price has to be lower than the selling price** as much as possible. The following prices are charged for camping articles in your shop:

back-pack, type A:	**fl 150,--**	**a piece**
sleeping-bag, type A:	**fl 150,--**	**a piece**
tent, type C:	fl 250,--	a piece
pans, type C:	fl 50,--	a set

Of course, you do not tell your friend these prices!

After the negotiations, you and your friend will have some beer, and you will talk about the performances of the sports club of which you are both members.

The role play has to last **at least half an hour**. Half an hour is a long time. You should therefore take your time for the amicable part of the role play. You can also take your time for the negotiations.

The purpose of this research is **not** to investigate selling techniques, argumentation structures, or your inventiveness. Therefore, do not pay much attention to what you say and how you talk: be/play yourself as much as possible, and talk the way you normally talk to friends.

C.2 Instructions for the salesman

You are going to visit a friend. The last time you saw him was at a party some days ago. You met many old friends at that party, and heard many fierce discussions (for instance about euthanasia, the education policy, religion). You and your friend will talk some time about that **party** (the people, the discussions, the food).

Then, you will ask your friend after his **adventures at the dentist**. He told you at the party that his wisdom-tooth had to be extracted. You want to know how that went. Moreover, you yourself have some stories about dentists, too.

Finally, you will talk about what you have been doing since the party (amusement parks, holidays, visits to family/friends, TV-programs, pets, cinema?).

After you have talked about several topics for minimally a quarter of an hour, you will tell your friend that you have not come only for pleasure. You want to sell him some products.

You are the salesman of **Campex**. Campex is a firm producing camping articles. It is famous for the high quality of its products.

Your friend is the owner of the shop **Sports & Travel**. Sports & Travel sells all kinds of articles in the area of sports and leisure time. Camping articles also belong to the assortment.

You want to sell to Sports & Travel **for prices as high as possible**:

- 100 back-packs, type A
- 100 sleeping-bags, type A
- 100 tents, type C

You want to sell these 300 articles as one unit. You have sold these articles to your friend before. There will therefore probably be no discussion about the exact types of the products.

You are allowed to ask different prices for the articles. When fixing the prices, you have to keep in mind the production prices. **The selling prices have to be higher than these production prices**. Back-packs, sleeping-bags, and tents have the following production prices:

back-pack, type A:	fl 50,--	a piece
sleeping-bag, type A:	fl 75,--	a piece
tent, type C:	fl 100,--	a piece

Of course, you do not tell your friend these prices!

After the negotiations, you and your friend will have some beer, and you will talk about the performances of the sports club of which you are both members.

The role play has to last **at least half an hour**. Half an hour is a long time. You should therefore take your time for the amicable part of the role play. You can also take your time for the negotiations.

The purpose of this research is **not** to investigate selling techniques, argumentation structures, or your inventiveness. Therefore, do not pay much attention to what you say and how you talk: be/play yourself as much as possible, and talk the way you normally talk to friends.

D Fragments of the corpus

The sentences are translated freely into English.

D.1 Free conversation between Subjects C and D in the author's presence

Subject D: *Nou, volgens mij was het idee dat daar een uitbater zou komen*
Well, according to me, the idea was that a business would come there

die wat gezelligheid zou brengen. Maar dat is nog niet gelukt.
which would liven things up a bit. But that hasn't happened yet.

Dat vind ik ook, die ...
I also find that, those...

Subject C: *Dat wist ik helemaal niet. Wat zouden die daar gaan doen?*
I didn't know that at all. What were they going to do there?

Patat uitventen?
Hawk chips?

Subject D: *Ja, iets met parasolletjes enzo en...*
Yes, something with little parasols and so on and...

Subject C: *Iets met parasolletjes?*
Something with little parasols?

Subject D: *Ja, iets met parasolletjes.*
Yes, something with little parasols.

Subject C: *Iets met parasolletjes.*
Something with little parasols.

Zo stond het ook in het voorstel waarschijnlijk.
That's probably also the way it was formulated in the proposal.

Subject D: *Ja, ik denk het wel.*
Yes, I think so.

Subject C: *Bestemmingsplan: iets met parasolletjes daar.*
Development plan: something with little parasols there.

Subject D: *Ja, en nu hebben ze van die leuke betonnen tafels neergezet.*
Yes, and now they have set up those nice concrete tables.

Is dat niet even...?
Isn't that just...?

Subject C: *Van die slachttafels?*
Those butcher's tables?

Subject D: *Ja precies.*
Yes precisely.

Subject C: *Van die islamitische slachttafels.*
Those Islamic butcher's tables.

Subject D: *Ja, die kan je goed schoonspuiten ja.*
Yes, they are easy to hose down yes.

D.2 Free conversation between Subjects K and L in the author's presence

Subject K: *Toen kenden we elkaar al wat beter.*
We knew each other a bit better by that stage.

Subject L: *Want ik zat vaak te werken bij Nederlands.*
Because I was often working at [the department of; ME] Dutch.

Ik was bij ATW.
I was in [the department of; ME] General Linguistics.

Maar ik vond de bibliotheek bij Nederlands prettiger zitten.
But the Dutch library was more comfortable.

En ik was met Nederlands bezig, met Middelnederlands vaak ook.
And I was working on Dutch, also often on Middle-Dutch.

Dus daar zat ik vaak, en eh hij is dan wel zo herderlijk dat ie...
So I was often there, and erm he is then so pastoral that he...

Subject K: *Begin je weer?*
Are you at it again?

Subject L: *Ja, ik begin weer, ja, ja, ja.*
Yes, I am at it again, yes, yes, yes.

Subject K: *Dat heb ik al weer een paar jaar niet,*
I haven't (heard) that for a few years,

heb ik al weer een paar jaar niet gehoord!
I haven't heard for a few years!

Subject L: *Ja, ik noem hem wel eens een pastoraal werker*
Yes, I sometimes call him a pastor

onder de taalkundigen.
among the linguists.

Subject K: *Dat bedoelde ik dus, ja.*
That's what I meant, yes.

ME: *Nou, dat is toch niet slecht.*
Well, that's surely not bad.

Subject L: *Ja, dat vindt hij wel.*
Yes, but he thinks so.

Subject K: *Een dergelijke rol wil ik niet hebben.*
I don't want to have such a role.

Subject L: *Nou je doet ook alles om die rol af te stoten.*
Well, you're doing everything to fight off that role.

Subject K: *Dank je, dat was ook weer niet de bedoeling.*
Thank you, I did not want to put it like that either.

Subject L: *Nee, maar dat heb je toch minder dan vroeger, merk ik.*
No, but you are less like that than in the past, I notice.

Subject K: *Nou, dat komt natuurlijk ook voor een deel doordat*
Well, that is naturally also partly due to the fact that

ik het drukker heb, want ik heb natuurlijk een heel ander soort baan.
I am busier, because I have naturally a very different type of job.

Subject L: *Ja, ja, je hebt het ontzettend druk gekregen de laatste tijd.*
Yes, yes, you have become awfully busy lately.

D.3 Free conversation between Subjects A and B during the role-play

Subject A: *Het waren dus schrijvers van reguliere romans en zo,*
So they were authors of regular novels and the like,

en daar de wat meer erotische fragmenten uit. Mensen als
and from those the somewhat more erotic fragments. People like

die franse schrijfster B.G., van "Zout op mijn huid", J.Z, M.D....
that French author B.G., of "Salt on my skin", J.Z, M.D....

Subject B: *In het GROTE muziekcentrum, in de grote zaal?*
In the LARGE music centre, in the big hall?

Subject A: *Nee, het was in de kleine zaal. Ben je daar wel eens geweest?*
No, it was in the small hall. Have you ever been there?

Subject B: *Oh, ik wou al zeggen. Dat is toch...*
Oh, I was going to say. That's surely...

Dat is toch geen sfeer om dat soort dingen te gaan...
That's surely no atmosphere ... such things...

Subject A: *Nee dat is gigantisch. Daar wordt wel het, eh hoe heet die ook weer?*
No that's gigantic. There the, erm what's that called again?

De nacht van de poezie, en zo,
The poetry night, and the like,

wordt wel altijd in de grote zaal gehouden.
is always organized in the big hall.

Subject B: *Ja, oké, maar dat heeft zijn eigen. Dat was ook wel eens*
Yes, all right, but that has its own. That was also once

in Carré en ook wel in Rotterdam wel eens geweest.
in Carré and also once in Rotterdam.

Dat heeft zijn eigen sfeer. Ik ben d'r nooit geweest,
That has its own atmosphere. I've never been there,

maar wel wat gezien op de televisie. En eh...
but have seen something on TV. And erm...

Subject A: *Nee, dit is in de kleine zaal, en dat is redelijk compact en...*
No, this is in the small hall, and that's fairly compact and...

Subject B: *En dat heeft ook een zeker theatraal effect.*
And that also has a certain theatrical effect.

Subject A: *Ja, dan zit je echt allemaal dicht om het podium voor.*
Yes, then you are really all sitting close to the front stage.

Subject B: *Als ik denk aan een schrijver die voordraagt*
When I think of an author who is reciting

of voorleest uit eigen werk, dan heb je toch een meer,
or reading aloud from his own work, then you would rather a more,

dan stel je je iets huiselijkers voor.
then you would imagine something more homely.

Subject A: *Ja, je kan je altijd nog betere locaties dan dit voor stellen.*
Yes, you could always imagine locations even better than this one.

D.4 Free conversation between Subjects M and N during the role-play

Subject M: *Dus ik heb ook inderdaad wel eens gedacht dat ik*
So indeed I also sometimes thought that I

het wel leuk zou vinden om af en toe eens op te passen op kinderen.
would like to take care of children every now and then.

Maar ik heb eigenlijk niemand die in de buurt woont.
But nobody is living in my neighbourhood.

Dus het komt er niet zo van.
So it does not really happen.

Subject N: *Ja, het is, het is gewoon leuk om ze te zien opgroeien.*
Yes, it's, it's just nice to see them grow up.

Het gaat ook zo snel, die veranderingen.
It's going so fast, those changes.

Subject M: *En doe je dat dan voor iemand die die wat moet doen*
And do you do that then for someone who who has to do something

op het gebied van werk ,of iemand die het leuk vindt om
with respect to work, or someone who likes to

eens een dag iets zonder kinderen te doen?
spend a day sometimes without children?

Subject N: *Eh, nou, ja, het is zo dat eh nou ze werken nou allebei.*
Erm, well, yes, it's the case that erm well they both work now.

Ze zijn. Hij heeft een zetterij. En die is dan achter het huis,
They are. He has a composing shop. And it is behind the house,

vlakbij. Dus d'r is wel eens een avond dat ze allebei gaan werken.
very near. So they are evenings they are both working.

Maar ze hebben ook zoiets van "we willen
But they also think "the two of us want

een avond per week iets met z'n tweeën gaan doen".
to do something together one evening a week".

En ik kom dus ook een vaste avond. En dan gaan ze uit,
And therefore I come on one fixed evening. And then they go out,

of uiteten, of wat dan ook. En ja, het is ook.
or have dinner somewhere, or whatever. And yes, it's also.

Ik ken die kinderen... Het duurt nu al nou 3,5 jaar.
I have known these children... It has now been going on for 3.5 years.

Ik begon er met eentje.
I started with one.

Subject M: *Dus je kent ze gewoon goed ook natuurlijk.*
So you know them very well of course.

Subject N: *Ja precies. Ik begon er met eentje van 1/2 jaar,*
Yes precisely. I started with one of 1/2 year old,

en inmiddels zijn het er drie. Dus ja.
and by now there are three of them. So yes.

Subject M: *Maar het is niet alleen maar 's avonds dan, want 's avonds*
But then it isn't only in the evening, because in the evening

liggen ze ook veel op bed, natuurlijk, als ze zo jong zijn.
they are mostly in their beds, of course, when they are so young.

Subject N: *Nou, het is het begin van de avond.*
Well, it's the beginning of the evening.

En ik stop ze dan in bed inderdaad.
And I do put them to bed.

Subject M: *Nou je hebt wel kans*
Well, you have the opportunity

om er wat wat mee te doen, in ieder geval.
to do something something with them, in any case.

Subject N: *Je doet er wat mee, je stopt ze in bed,*
You do something with them, you put them to bed,

en voor de rest heb je best een rustig avondje.
and for the rest you've got a fairly quiet evening.

D.5 Negotiations between Subjects F and G

Subject G: *Ik zei wel "ze kosten mij dan 100 gulden eh het stuk",*
I did say "they've cost me then 100 guilders erm a piece",

maar daar kan natuurlijk wel,
but of course it is possible,

dat je wel al 10, 15 gulden af kan krijgen, als je dan ...
that you can lower it by at least 10, 15 guilders, if you then...

Subject F: *Ja, nou ja, dan, dan wordt het voor mij ook weer wat*
Yes, well yes, in that case, it becomes again also for me a bit

interessanter om om dat te proberen. Want op zich hebben ze ook wel
more interesting to to try that. Because I agree it's

kwaliteit hoor. Ik, ik. Het is niet hè. Want die andere rits is
quality stuff. I, I. It is not, isn't it. Because that other zipper is

bijvoorbeeld heel handig, dat als je even die rits
for instance very convenient, that if you just

in z'n geheel los maakt, dan kun je hem groter maken, hè.
unzip that zipper completely, you can then make it larger, can't you.

Subject G: *Het is een uitstekend...*
It's an excellent...

Subject F: *En dat, daarvan zeggen de mensen "dat is heel handig",*
And that, people are always saying of it "that's very convenient",

want je gaat zo'n berg op met voorraden, hè,
because you're climbing such a mountain with supplies, don't you,

in zo'n trektocht. En al al onderweg drink je dingen op, en eet je
in such a hike. And on the way you start to drink things, and eat

dingen op, en dan, dan kun je hem later wat kleiner maken,
things, and then, then you can make it a bit smaller later on,

en dat is handig met....Dus dat is een heel handig element d'r in.
and that's convenient with... So that's a very convenient element in it.

Dus dat vind ik heel goed, heel inventief op zich.
So I think that's truly very good, very inventive.

Subject G: *Ja, ik zit nu te denken, ik denk dat je toch wel zo'n 100 stuks...*
Yes, I'm now thinking, I think that you ... at least 100 pieces...

Die kan ik wel goedkoop krijgen dan. Ik denk dat je dan
I can get them cheaply then. I think that you then

voor een, ja, 80 gulden het stuk een eerste kwaliteit rugzak hebt.
have a first quality back-pack for, yes, 80 guilders a piece.

En dan, voor de, toch ook voor het voorjaar en de zomer.
And then, for the, surely also for the spring and the summer.

D.6 Negotiations between Subjects O and P

Subject P: *Nou, van ten, van tenten zowiezo, daar heb ik net eh*
Well, of ten, of tents, in any case, I just ... some erm

van een andere... Een andere aanbieder die kwam langs,
from another... Another salesman just called,

en dat waren prachtige tenten voor een prijs...
and those were beautiful tents for a price...

Subject O: *Maar T. toch, T. toch! Jij hebt vorig jaar nog gezegd van*
But T., T.! You said only last year

"kom met zo'n zelfde lading weer terug". Dus ik heel m'n best doen
"Bring the same stuff next time". So, there am I trying my very best

om hier die tenten natuurlijk los te peuteren. En wat ga je
to scrape those tents together for you. And what do you

me nou vertellen? Dat je bij een ander de rotzooi vandaan haalt!
tell me now? That you get the caboodle from somebody else!

Het is toch niet hetzelfde hè? Ik bedoel, het is toch niet dezelfde soort
It isn't the same, isn't it? I mean, it isn't the same type of

eh tent? Ik bedoel, zo'n kwaliteit die krijg je bij een ander niet.
erm tent? I mean, this quality you don't get from someone else.

Subject P: *Nee, het is een, het is een ander soort tent.*
No, it's a, it's another type of tent.

Subject O: *Dat wou ik zeggen.*
That's what I wanted to say.

Subject P: *Alleen, ja, hij had ze wat eerder.*
Just, yes, he had them a bit earlier.

Subject O: *Ja oké.*
Yes all right.

Subject P: *Toen ik ze nodig had.*
When I needed them.

Subject O: *Oké, nee oké, ja.*
All right, no all right, yes.

Subject P: *Dat is het probleem. Ik had ze nodig, nou toen heb ik ze gekregen.*
That's the problem. I needed them, well then I got them.

Subject O: *Ja oké, maar je had me even moeten bellen joh.*
Yes all right, but you should just have called me, mate.

Waarom heb je dan niet gebeld? Dat is natuurlijk wel weer even...
Why didn't you call? That's of course again just...

Subject P: *Ja, maar ik heb je gebeld, maar je secretaresse zei dat je was of in*
Yes, but I did call you, but your secretary said that you were either in

Maastricht of in Groningen. En dan zeg ik "bel me eens terug".
Maastricht or in Groningen. And then I say "call me back some time".

Subject O: *En dan ... ja, ik was voor...*
And then ... yes, I was for...

Subject P: *En dan belde jij terug als ik weg was.*
And then you called back when I was away.

Dus dat ging ook nooit...
So that also never went...

Subject O: *Ik was voor jou...*
I was for you...

Ik was voor jou aan het onderhandelen natuurlijk.
I was negotiating for you of course.

E Judgements of the recorded speech

E.1 Ages and origins of the judges

Table E.1 Ages and origins of the judges.

Judge	Age	Origin
1	38	Western part of the Netherlands
2	37	Western part of the Netherlands
3	28	Western part of the Netherlands
4	47	Eastern part of the Netherlands
5	42	South-Eastern part of the Netherlands
6	28	South-Western part of the Netherlands
ME	27	Western part of the Netherlands

E.2 The judgements

The judgements were originally in Dutch.

- Judgments of Subject A's speech

Judge 1:	Standard Dutch, sometimes sloppy.
Judge 2:	Standard Dutch, but perhaps not completely Standard Dutch with respect to the vowels.
Judge 3:	Standard Dutch.
Judge 4:	Standard Dutch.
Judge 5:	Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject B's speech

Judge 1:	Standard Dutch, well-articulated.
Judge 2:	Neat Standard Dutch.
Judge 3:	Standard Dutch.
Judge 4:	Standard Dutch.
Judge 5:	Very educated Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject C's speech

Judge 1:	Standard Dutch with a flavour of the dialect of Amsterdam.
Judge 2:	Standard Dutch.
Judge 3:	Standard Dutch, well formulated.
Judge 4:	Standard Dutch.
Judge 5:	A western variant of Standard Dutch.
Judge 6:	Standard Dutch, a bit posh.
ME:	Standard Dutch.

- Judgements of Subject D's speech

Judge 1:	Standard Dutch with a flavour of the dialect of Amsterdam.
Judge 2:	Standard Dutch.
Judge 3:	Standard Dutch, well formulated.
Judge 4:	Standard Dutch with a flavour of the dialect of Amsterdam.
Judge 5:	Standard Dutch.
Judge 6:	Standard Dutch with a flavour of the dialect of Amsterdam.
ME:	Standard Dutch with a flavour of the dialect of Amsterdam.

- Judgements of Subject E's speech

Judge 1:	Standard Dutch.
Judge 2:	A western variant of Standard Dutch.
Judge 3:	Standard Dutch with a flavour of the dialect of Amsterdam.
Judge 4:	Standard Dutch with a flavour of the dialect of Amsterdam.
Judge 5:	A western variant of Standard Dutch.
Judge 6:	Standard Dutch with a flavour of the dialect of Amsterdam.
ME:	Standard Dutch with a flavour of the dialect of Amsterdam.

- Judgements of Subject F's speech

Judge 1:	Distinguished Standard Dutch.
Judge 2:	Standard Dutch which is sometimes very polished.
Judge 3:	Distinguished, educated Standard Dutch.
Judge 4:	Distinguished Standard Dutch.
Judge 5:	Educated Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject G's speech

Judge 1:	Standard Dutch.
Judge 2:	Standard Dutch.
Judge 3:	Standard Dutch, well formulated.
Judge 4:	Standard Dutch.
Judge 5:	Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject H's speech

Judge 1:	Fairly Standard Dutch.
Judge 2:	Educated Standard Dutch.
Judge 3:	Standard Dutch.
Judge 4:	Standard Dutch.
Judge 5:	Standard Dutch.
Judge 6:	Standard Dutch with a flavour of the dialect of South-Holland, a bit posh.
ME:	Standard Dutch.

- Judgements of Subject I's speech

Judge 1:	Standard Dutch.
Judge 2:	Very beautiful Standard Dutch.
Judge 3:	Standard Dutch, sometimes sloppy.
Judge 4:	Standard Dutch.
Judge 5:	Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch with a slight flavour of the dialect of Amsterdam.

- Judgements of Subject J's speech

Judge 1:	Not completely Standard Dutch.
Judge 2:	Educated Standard Dutch.
Judge 3:	Standard Dutch, well formulated.
Judge 4:	Standard Dutch.
Judge 5:	Educated Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject K's speech

Judge 1:	Fairly Standard Dutch.
Judge 2:	Educated Standard Dutch.
Judge 3:	Distinguished Standard Dutch.
Judge 4:	Standard Dutch.
Judge 5:	Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject L's speech

Judge 1:	Fairly Standard Dutch.
Judge 2:	Polished Standard Dutch.
Judge 3:	Standard Dutch.
Judge 4:	Standard Dutch.
Judge 5:	A western variant of Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject M's speech

Judge 1:	Approximately Standard Dutch, sloppy.
Judge 2:	Standard Dutch.
Judge 3:	Standard Dutch with a flavour of the dialect of North-Holland.
Judge 4:	Standard Dutch with a flavour of the dialect of Amsterdam.
Judge 5:	Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch with a slight flavour of the dialect of North-Holland.

- Judgements of Subject N's speech

Judge 1:	Standard Dutch.
Judge 2:	Standard Dutch.
Judge 3:	Standard Dutch.
Judge 4:	Standard Dutch.
Judge 5:	Approximately Standard Dutch. Large influence of the dialect of Amsterdam.
Judge 6:	Standard Dutch.
ME:	Standard Dutch with a slight flavour of the dialect of Amsterdam.

- Judgements of Subject O's speech

Judge 1:	Standard Dutch with a flavour of the dialect of Amsterdam.
Judge 2:	Standard Dutch.
Judge 3:	Standard Dutch.
Judge 4:	Standard Dutch.
Judge 5:	A western variant of Standard Dutch.
Judge 6:	A western variant of Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject P's speech

Judge 1:	Standard Dutch, a bit posh.
Judge 2:	Standard Dutch.
Judge 3:	Distinguished Standard Dutch.
Judge 4:	Educated Standard Dutch, carefully articulated.
Judge 5:	A western variant of Standard Dutch.
Judge 6:	Standard Dutch.
ME:	Standard Dutch.

- Judgements of Subject Q's speech

Judge 1:	A western variant of Standard Dutch.
Judge 2:	Strange vowels.
Judge 3:	Dialect from the city of Zaanstad.
Judge 4:	Strange vowels.
Judge 5:	Not Standard Dutch.
Judge 6:	A western variant of substandard Dutch, sloppy.
ME:	A western variant of substandard Dutch.

F Separating two clusters in two-dimensional plots

F.1 Introduction

Chapters 8 and 9 presented figures in which the closure durations of voiced and voiceless alveolar stops are plotted against their burst durations. The voiced and voiceless stops appeared to be grouped in clusters, that is, there is an area of voiced stops and an area of voiceless stops in each plot. The boundaries between the areas were computed as well as the accuracy with which these boundaries separate the voiced and voiceless stops.

Two existing methods with which the boundaries could have been computed are Linear Discriminant Analysis and Classification And Regression Trees (Toni Rietveld, personal communication). We did not use these methods in the present study, because they do not allow us to determine whether the differences between computed boundaries for different plots are statistically different, and this was necessary in section 9.2.

We used an alternative approach which has been developed by Evert Wattel and ourselves, and is described in this appendix. By this method the areas of voiced and voiceless stops in a plot are separated as well as possible by a straight separation line (§F.2), which is characterized by its position and slope. The position of a separation line is defined as the length of the line which is perpendicular to that separation line and connects it with the origin of the coordinate system. The slope of a separation line is the angle of this perpendicular line with the X-axis (see Figure F.1 on the next page). Standard errors can be assigned to the computed position and slope of each separation line (§F.3), and on the basis of these standard errors it can be determined whether there is a statistically significant difference between two lines (§F.4). The method separates areas of voiced and voiceless stops at least as well as Linear Discriminant Analysis and Classification And Regression Trees (§F.5).

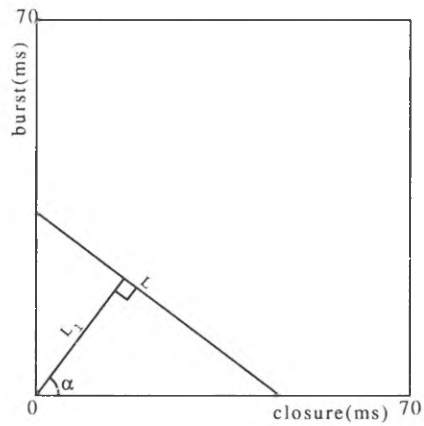


Figure F.1 Position (L_1) and slope (α) of separation line L.

F.2 Computation of the separation line

The calculation of the optimal separation line consists of the following steps.

1. Calculating the position and slope of the line which connects the average closure and burst duration of the voiced stops (henceforth *d*'s) to the average closure and burst duration of the voiceless stops (henceforth *t*'s).
2. Calculating the position and slope of the line which bisects the former line perpendicularly and exactly in the middle. This latter line is a first estimation of the separation line.
3. Finding the slope of the optimal separation line by changing the slope of the line obtained in (2) in very small steps until the error of the line increases.
4. Finding the position of the optimal separation line by changing the position of the line obtained in (3) in very small steps until the error of the line increases.
5. The resulting line is the optimal line.

The error of a separation line is related to the positions of the voiced and voiceless stops in the plot: the more stops are positioned on the wrong side of the line, i.e. are positioned in the wrong area, the greater the error. Since the closure and burst durations of each stop, and therefore its position in a plot, were determined with a certain variance due to measurement error (§8.3.5.2), some stops are perhaps positioned on the wrong side of the line, whereas they actually belong to its correct side, and vice versa. The further a stop is removed from the line, the smaller the probability that this is the case. Hence, stops on the wrong side of the line should contribute less to the error if they are nearer the line, and stops on the correct side of the line should contribute less if they are further removed. The error of a line is therefore also related to the distance between the line and the stops.

The error is computed as follows. Each stop is assigned an error factor, which is related to the side of the line on which it is located as well as to its distance to the line. Figure F.2 plots the error factor as a function of the distance for stops located on the wrong side of the line, while Figure F.3 plots the error factors for stops on the correct side.

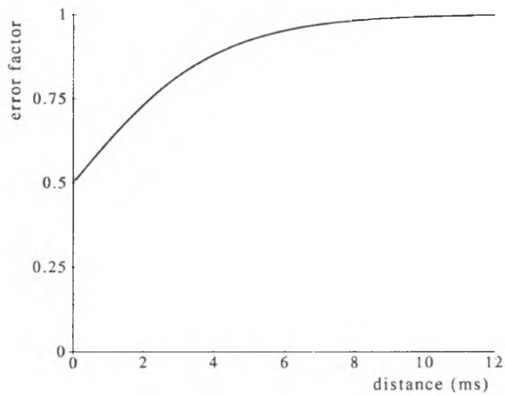


Figure F.2 The error factor for stops positioned on the wrong side of the separation line as a function of their distance to the line.

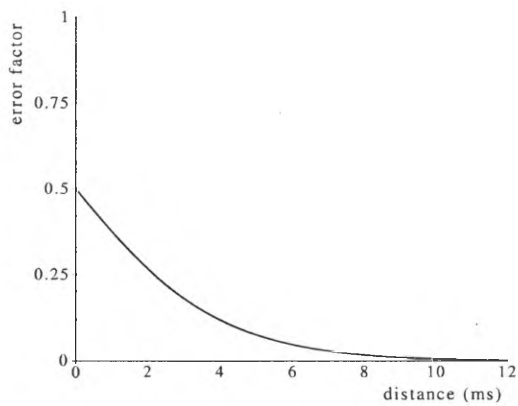


Figure F.3 The error factor for stops positioned on the correct side of the separation line as a function of their distance to the line.

The error factor of each voiced stop is multiplied by the number of *t*'s in the plot, while the error factor of a voiceless stop is multiplied by the number of *d*'s in the plot. The sum of the multiplication is the error of the line.

The reason for multiplying the error factor of a stop by the number of stops of the opposite classification is the following. Imagine a plot which contains more *d*'s than *t*'s, and that a new *t* and a new *d* are added to this plot. The probability that the new *t* is located far from the other *t*'s is greater than the probability that the new *d* is far removed from the other *d*'s. The line which best separates the *t*-area and *d*-area in such a plot is therefore further removed from the *t*'s than from the *d*'s, which is the case if the error of each *t* is relatively greater than the error of each *d*. This is obtained by multiplying the error factor of each stop by the number of stops of the opposite classification in the plot.

Figures F.2 and F.3 show that all stops which are located at a distance of minimally 6 ms on the same side of the line are assigned approximately the same error factor. No difference is made between, for instance, a stop located at a distance of 7 ms and a stop at a distance of 9 ms. The reason for this is that stops which are minimally 6 ms removed from the line are assumed not to be positioned on that particular side of the line because of measurement errors, and therefore count as (nearly) fully wrong or correct. This assumption is based on the observation that only a few stops which are at a minimal distance of 6 ms from the line are positioned on the wrong side of the line according to their auditory classifications, and on the results of the double measurements of 100 closure and burst durations described in section 8.3.5.2, which show that only few stops may be expected to be located further than 6 ms from their proper position in the plot.

F.3 Standard errors of the position and slope of the line

The standard errors of the position and slope of a separation line may be assumed merely to indicate the adequacy of the separation, i.e. the proportion of stops which is positioned on the wrong side of the line according to their auditory classifications. In that case, the errors increase when the amount of data increases, as the more data are available, the more difficult it is to separate the *t*'s and the *d*'s. This is counter-intuitive, since the characteristics of the line are expected to be more valid when they are based on more data instead of fewer ones.

The standard errors should indicate the probability that the position and slope of a line change when a new stop is incorporated into the plot. This is the case if the characteristics of the line are computed as many times as there are stops in the corpus, with every computation being based on all stops in the data minus one, and every stop being removed from the data set once, and the standard errors are defined as in (1).

(1) Standard error A (S_A) $$= \frac{\sqrt{(n-1)} * \sqrt{\sum_{i=1}^{n} (A_i - \bar{A})^2}}{\sqrt{n}}$$

A = the slope (ms per ms) or position (ms) of the line.
Ā = the average value of A.

The standard errors then decrease when the number of data increases, since the more stops are present in the figure, the less the absence of one them affects the characteristics of the line. They therefore adequately indicate the validity of the line.

If the standard errors of the position and slope of the optimal separation line are computed with formula (1), they are only based on those stops that were unanimously classified as voiced or as voiceless. These stops form 84% of the total number of alveolar stops judged in the present study. The standard errors, however, should also be based on the stops that were not unanimously classified, and are not incorporated in the final data set, for the following reason. The *t*'s and *d*'s can be assumed to be positioned in the figures according to Gaussian curves: the further a position is removed from the centre of the cluster of the *t*'s, the smaller its chance to accommodate a *t*, and the further it is removed from the centre of the cluster of the *d's* the smaller its chance to accommodate a *d*. The 16% of stops that were not unanimously classified as either voiced or voiceless are presumably located at positions with small probabilities to accommodate a *d* or *t* respectively: they form the tails of the Gaussian curves. If these stops are not taken into account, the Gaussian curves lack 8% of each of their tails, and the computed standard errors are smaller than the actual ones.

The ratio between the computed standard errors and the actual standard errors, then, can be computed as follows. Under the assumption that all *t*'s, i.e. all stops which should have been classified as voiceless, are distributed in the plots according to a Gaussian distribution, these *t's* have a distribution with variance (2). Similarly, all *d*'s have a distribution with variance (2).

(2) $$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2\sigma^2}} dx = \sigma^2$$

Under the assumption that the stops which are not unanimously classified as voiced or voiceless form the tails of the Gaussian curves, the variance of the Gaussian distributions for the *t*'s or *d*'s incorporated in the figures is represented by formula (3), with *T* and -*T* indicating the positions where the tails of the Gaussian curves are cut off.

(3) $$\frac{1}{\sigma\sqrt{2\pi}}\int_{-T}^{T} x^2 e^{-\frac{x^2}{2\sigma^2}}dx$$

The variable *T* can be determined by means of equation (4). The formula on the left side of this equation represents the surface under the Gaussian curve for the *t*'s or the *d's* in the figures. It equals 0.84, since these figures represent 84% of the stops.

(4) $$\frac{1}{\sigma\sqrt{2\pi}}\int_{-T}^{T} e^{-\frac{x^2}{2\sigma^2}}dx = 0.84$$

T appears to be 1.4, and when this value is substituted for *T* in formula (3), the variance of the Gaussian curve for the *t's* or the *d's* in the plots appears to be ½ of the variance of the Gaussian curve for all *t*'s and *d*'s. Since the standard error is the square root of the variance, this means that the standard error of the stops in the plot is the square root of ½ times the standard error of all stops, and that the actual standard error of the position or slope of a separation line is (5).

(5) standard error A (S_A) $$= \sqrt{(n-1)} * \frac{\sqrt{\sum_{i=1}^{n}(A_i-\bar{A})^2}}{\sqrt{n}} * \frac{1}{\sqrt{2}}$$

A = the slope (ms per ms) or position (ms) of the line.
Ā = the average value of A.

F.4 Deciding whether lines are different

If the positions and slopes, as well as the standard errors of these characteristics of two lines, are known, it is possible to determine whether the difference between these lines is statistically significant. The positions of the lines should be plotted against their slopes, and the resulting points be surrounded by ellipses indicating the confidence intervals of the characteristics. If the ellipses are well apart, the lines can be assumed to be different with a certainty corresponding to the confidence interval (see Figure F.4 for an example).

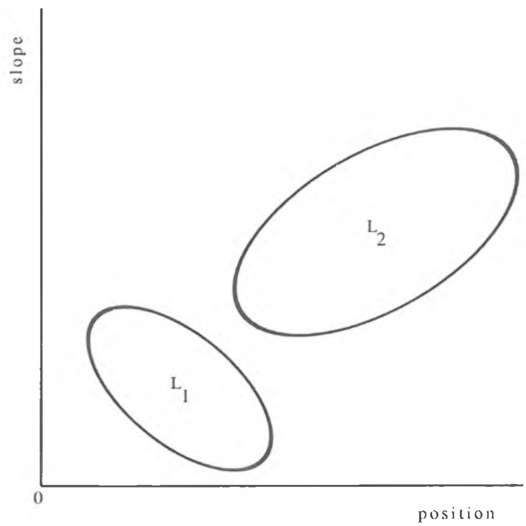


Figure F.4 The positions and slopes of lines L_1 and L_2, which differ significantly ($p < 0.05$).

For $\alpha = 0.05$, the radii of the ellipse in the horizontal and vertical directions are those in (6).

(6) a. Radius in the horizontal direction $= S_{position} * \frac{1}{2}\sqrt{6}$

b. Radius int the vertical direction $= S_{slope} * \frac{1}{2}\sqrt{6}$

Formula (6) was determined as follows. It is assumed that the closure and burst duration of every stop have a standard normal distribution. The average closure duration and the average burst duration of the stops, then, have normal distributions as well, as have the points which represent the position and slope of the separation lines in plots. Since a normal distribution with two variables (e.g. position and slope) has an α of 0.05 at a position of $6^{½}\sigma$, with σ representing the standard error of one variable in the horizontal direction and the standard error of the other variable in the vertical direction, the probability is minimally 95% that a point in a figure that is outside an ellipse with a radius of $6^{½}\sigma$, with σ representing the standard error of the position or the slope of the relevant line, represents another line than the centre of the ellipse does. This means that the probability is minimally 5% that the centres of two non-overlapping ellipses with radii of ½ $6^{½}\sigma$ do not represent different lines.

F.5 Comparison with alternative approaches

As mentioned in section F.1, there are two alternatives to the approach We have chosen here for the determination of the boundaries between the areas of the *t*'s and the areas of the *d*'s in the figures: Linear Discriminant Analysis (Goldstein & Dillon 1978) and Classification And Regression Trees (see e.g. Breiman et al. 1984; Steinberg & Colla 1995). In the remainder of this appendix, these two methods will be referred to as LDA and CART respectively. As a validation of the method that we have chosen here (henceforth method Wattel & Ernestus, or W&E for short), Tables F.1 to F.3 list the percentages of *t*'s and *d*'s that are contained in the areas of *t*'s and *d*'s in three plots when these areas are computed with the three methods. The relevant plots are those for which separation lines were calculated in this study: the plot containing all alveolar stops (Figure 8.7 in Chapter 8), the plot of stops following [+high] vowels (Figure 9.1 in Chapter 9), and the plot of stops following [-low] vowels (Figure 9.2 in Chapter 9). The results produced by LDA and CART were provided by Toni Rietveld. Inspection of the three tables show that the methods assign approximately the same percentages of stops to the corresponding areas, which means that the methods have apparently the same adequacy.

Table F.1 The relative (%) numbers of d*'s and* t*'s of the complete data set (341* t*'s and 308* d*'s) in the corresponding areas of* d*'s and* t*'s computed with LDA, CART, and W&E.*

Method	Percentage of	
	t's in the *t*-area	*d*'s in the *d*-area
LDA	91%	94%
CART	93%	92%
W&E	93%	94%

Table F.2 The relative (%) numbers of d*'s and* t*'s following [+high] vowels (106* t*'s and 208* d*'s) in the corresponding areas of* d*'s and* t*'s computed with LDA, CART, and W&E.*

Method	Percentage of	
	t's in the *t*-area	*d*'s in the *d*-area
LDA	91%	96%
CART	91%	90%
W&E	88%	95%

Table F.3 The relative (%) numbers of d's and t's following [-high] vowels (197 t*'s and 94* d*'s) in the corresponding areas of* d*'s and* t*'s computed with LDA, CART, and W&E.*

Method	Percentage of	
	t's in the *t*-area	*d*'s in the *d*-area
LDA	93%	91%
CART	95%	93%
W&E	98%	89%

G Classifications of the single intervocalic stops

General remarks:

- The positions of the main stresses are indicated both in content words and function words. These main stresses are not always realized.
- Unless indicated otherwise, the verb is in a singular form.
- The string "-s" denotes the verbal suffix third person singular present tense.
- The listed frequencies of occurrences are the numbers of occurrence in our corpus and in the pilot study of the Eindhoven corpus of spoken Dutch (Uit den Boogaart 1975).

Table G.1 Numbers of voiced and voiceless stops in tokens of verb form + ik. *The numbers are broken down by the place of articulation of the stop, and the type of verb form. The verb form types are ordered by their frequencies of occurrence in verb form +* ik, *which are listed as well.*

Place of articulation	Verb form			Frequency	Numbers of stops	
					voiced	voiceless
Alveolar	*weet*	/'ʋet/	'know'	148	39	9
	moet	/'mut/	'must'	102	36	8
	had	/'hɑd/	'had'	89	27	7
	laat	/'lat/	'let'	38	3	1
	zit	/'zɪt/	'sit'	24	10	2
	zat	/'zɑt/	'sat'	11	2	0
	deed	/'ded/	'did'	7	1	0
	zet	/'zɛt/	'set'	4	0	1
	bied	/'bid/	'offer'	3	2	0
	schat	/'sxɑt/	'estimate'	3	1	1
	eet	/'et/	'eat'	1	1	0
	leid	/'lɛid/	'conduct'	1	1	0
	meet	/'met/	'mesaure'	1	1	0
	mijd	/'mɛid/	'avoid'	1	0	1
	praat	/'prat/	'talk'	1	0	1
Bilabial	*heb*	/'hɛb/	'have'	354	11	4
	begrijp	/bə'xrɛip/	'understand	19	5	0
	snap	/'snɑp/	'understand	7	0	1
	koop	/'kop/	'buy'	5	3	1
	hoop	/'hop/	'hope'	5	0	1
	liep	/'lip/	'walked'	4	1	0
	loop	/'lop/	'walk'	3	1	0
	riep	/'rip/	'called'	1	0	1
	stap	/'stɑp/	'step'	1	1	0
	type	/'tip/	'type'	1	0	1

Table G.2 Numbers of voiced and voiceless stops in tokens of verb form + het. *The numbers are broken down by the place of articulation of the stop, and the type of verb form. The verb forms types are ordered by their frequencies of occurrence in verb form +* het, *which are listed as well.*

Place of articulation	Verb form			Frequency	Numbers of stops	
					voiced	voiceless
alveolar	*weet*	/ˈʋet(-t)/	‘know(s)’	59	12	4
	had	/ˈhɑd/	‘had’	36	1	6
	doet	/ˈdu-t/	‘does’	15	1	3
	deed	/ˈded/	‘did’	14	4	0
	moet	/ˈmut(-t)/	‘must’	13	5	1
	ziet	/ˈzi-t/	‘sees’	4	1	0
	bied	/ˈbid/	‘offer’	3	2	0
	gaat	/ˈxa-t/	‘goes’	3	1	0
	laat	/ˈlat/	‘leave’	2	0	1
	haat	/ˈhat/	‘hate’	1	0	1
	spuit	/ˈspœyt/	‘spirt’	1	1	0
bilabial	*heb*	/ˈhɛb/	‘have’	70	6	3
	begrijp	/bəˈxrɛip/	‘understand’	6	1	0
	snap	/ˈsnɑp/	‘understand’	3	0	1

Table G.3 Numbers of voiced and voiceless stops in tokens of verb form + er. *The numbers are broken down by the place of articulation of the stop, and the type of verb form. The verb form types are ordered by their frequencies of occurrence in verb form +* er, *which are listed as well.*

Place of articulation	Verb form			Frequency	Numbers of stops	
					voiced	voiceless
alveolar	*moet*	/'mut/	'must'	42	1	6
	gaat	/'xa-t/	'goes'	27	1	6
	zit	/'zɪt(-t)/	'sit(s)'	25	3	6
	had	/'hɑd/	'had'	20	0	1
	staat	/'sta-t/	'stands'	15	1	3
	zat	/'zɑt/	'sat'	8	0	4
	ziet	/'zi-t/	'sees'	7	0	4
	doet	/'du-t/	'does'	3	0	2
	zet	/'zɛt(-t)/	'put(s)'	2	0	3
	pleit	/'plɛit/	'plead'	1	0	1
bilabial	*heb*	/'hɛb/	'have'	45	2	9
	koop	/'kop/	'buy'	2	0	2
	liep	/'lip/	'walked'	1	0	1

Table G.4 Numbers of voiced and voiceless /t/s followed by tautomorphemic schwa, broken down by morpheme type.

Morpheme type			Numbers of stops	
			voiced	voiceless
buiten	/'bœytən/	'outside'	2	20
cassette	/kɑ'sɛtə/	'cassette'	0	1
categorie	/kɑtəxo'ri/	'category'	0	1
computer	/kɔm'pjutər/	'id'	0	10
feite	/'fɛitə/	'effect'	0	4
klote	/'klotə/	'bloody'	1	0
letter	/'lɛtər/	'letter'	0	14
literair	/litə'rɛ:r/	'literary'	0	6
literatuur	/litəra'tyr/	'literature'	0	3
mate	/'matə/	'extent'	0	3
materiaal	/matəri'al/	'material'	0	1
meter	/'metər/	'meter'	0	4
ontzettend	/ɔnt'zɛtənd/	'terribly'	4	2
roulette	/ru'lɛtə/	'id'	0	1
route	/'rutə/	'id'	0	1
sateliet	/satə'lit/	'satelite'	0	1
sleutel	/'sløtəl/	'key'	0	1
theater	/te'atər/	'id'	0	2
titel	/'titəl/	'title'	0	1
uiteraard	/œytə'rard/	'of course'	0	3
water	/'ʋatər/	'water'	0	15
zaterdag	/'zatərdɑx/	'Saturday'	1	4

Table G.5 Numbers of voiced and voiceless /t/s followed by the infinitive marker -en, *broken down by word type.*

Word type			Numbers of stops	
			voiced	voiceless
besluiten	/bəˈslœytən/	‘to decide’	0	1
eten	/ˈetən/	‘to eat’	0	5
fluiten	/ˈflœytən/	‘to whistle’	0	2
heten	/ˈhetən/	‘to be called’	0	3
laten	/ˈlatən/	‘to let’	0	5
letten	/ˈlɛtən/	‘to pay attention’	0	2
meten	/ˈmetən/	‘to measure’	0	2
moeten	/ˈmutən/	‘must’	2	10
opschieten	/ˈɔpsxitən/	‘to hurry up’	0	1
praten	/ˈpratən/	‘to talk’	1	15
schieten	/ˈsxitən/	‘to shoot’	0	2
slijten	/ˈslɛitən/	‘to wear out’	0	2
sluiten	/ˈslœytən/	‘to close’	0	6
smijten	/ˈsmɛitən/	‘to hurl’	0	1
spotten	/ˈspɔtən/	‘to mock’	0	1
stoten	/ˈstotən/	‘to hit’	0	1
vatten	/ˈvɑtən/	‘to grasp’	0	4
weten	/ˈʋetən/	‘to know’	1	7
zetten	/ˈzɛtən/	‘to put’	1	16
zitten	/ˈzɪtən/	‘to sit’	5	24

Table G.6 Numbers of voiced and voiceless /p/s followed by tautomorphemic schwa, broken down by morpheme type.

Morpheme type			Numbers of stops	
			voiced	voiceless
appel	/'ɑpəl/	'apple'	0	1
kapper	/'kɑpər/	'hair dresser'	0	1
knipper	/'knɪpər/	'flash'	0	1
koepel	/'kupəl/	'dome'	0	1
koper	/'kopər/	'copper'	0	1
koppel	/'kɔpəl/	'couple'	0	1
open	/'opən/	'open'	0	5
operatie	/opə'rati/	'surgery'	0	2
opereren	/opə'rerən/	'to perform surgery'	0	2
peper	/'pepər/	'pepper'	0	1
principe	/prɪn'sipə/	'principle'	1	7
stapel	/'stapəl/	'pile'	0	3
super	/'sypər/	'super'	0	6
trappel	/'trɑpəl/	'stamp one's feet'	0	1
type	/'tipə/	'type'	0	18

Table G.7 Numbers of voiced and voiceless /p/s followed by the infinitive morpheme -en, *broken down by word type.*

Word type			Numbers of stops	
			voiced	voiceless
begrijpen	/bə'xrɛipən/	'to understand'	0	1
diepen	/'dipən/	'to deepen'	0	2
grijpen	/'xrɛipən/	'to grasp'	0	1
hopen	/'hopən/	'to hope'	0	1
kloppen	/'klɔpən/	'to knock'	0	2
knippen	/'knɪpən/	'to cut'	0	2
kopen	/'kopən/	'to buy'	0	16
lopen	/'lopən/	'to walk'	0	10
slapen	/'slapən/	'to sleep'	0	2
slepen	/'slepən/	'to drag'	0	1
stappen	/'stɑpən/	'to step'	0	1
stoppen	/'stɔpən/	'to stop/put'	0	5
typen	/'tipən/	'to type'	0	1

Table G.8 Numbers of voiced and voiceless /d/s followed by tautomorphemic schwa, broken down by morpheme type.

Morpheme type			Numbers of stops	
			voiced	voiceless
adem	/'adəm/	'breath'	1	0
beide	/'bɛidə/	'both'	4	0
beneden	/be'nedən/	'down'	6	0
bodem	/'bodəm/	'bottom'	2	0
elektrode	/elɛk'trodə/	'electrode'	1	0
geleden	/gə'ledən/	'ago'	9	0
heden	/'hedən/	'today'	1	0
ieder	/'idər/	'every'	5	0
iedereen	/idə'ren/	'everybody'	5	0
inmiddels	/ɪ'mɪdəls/	'by now'	1	0
jongstleden	/'jɔŋstledən/	'last'	1	0
kader	/'kadər/	'frame'	1	0
ladder	/'lɑdər/	'ladder'	1	0
mede	/'medə/	'co-'	2	0
methode	/me'todə/	'method'	2	0
middel	/'mɪdəl/	'middle' or 'means'	7	0
midden	/'mɪdən/	'middle'	8	0
mode	/'modə/	'fashion'	3	0
moeder	/'mudər/	'mother'	6	0
nader	/'nadər/	'approach'	2	0
neder	/'nedər/	'down'	3	0
periode	/peri'odə/	'period'	7	0
reden	/'redən/	'reason'	6	0
roddel	/'rɔdəl/	'gossip'	1	0
salade	/sa'ladə/	'salad'	1	0
schade	/'sxadə/	'damage'	2	0
tevreden	/tə'vredən/	'satisfied'	6	0
vader	/'vadər/	'father'	5	0
vergader	/vɛr'xadər/	'assemble'	2	0
verleden	/vɛr'ledən/	'past'	5	0
wedde	/'ʋɛdə/	'pay'	1	0

Table G.9 Numbers of voiced and voiceless /d/s followed by the infinitive morpheme -en, *broken down by word type.*

Word type			Numbers of stops	
			voiced	voiceless
beïnvloeden	/bəˈɪnvludən/	'to influence'	1	0
bereiden	/bəˈrɛidən/	'to prepare'	1	0
besteden	/bəˈstedən/	'to spend'	1	0
bieden	/ˈbidən/	'to offer'	14	0
luiden	/ˈlœydən/	'to ring'	1	0
optreden	/ˈɔptredən/	'to action'	1	0
raden	/ˈradən/	'to recommend'	2	0
scheiden	/ˈsxɛidən/	'to separate'	1	0
schudden	/ˈsxʏdən/	'to shake'	2	0
uitbreiden	/ˈœytbrɛidən/	'to extend'	2	0
waden	/ˈʋadən/	'to wade'	1	0

Table G.10 Numbers of voiced and voiceless /b/s followed by tautomorphemic schwa or the infinitive marker -en, *broken down by word type.*

Word type			Numbers of stops	
			voiced	voiceless
acceptabel	/ɑksɛp'tabəl/	'acceptable'	1	0
belabberd	/bə'lɑbərd/	'rotten'	0	1
blubber	/'blʏbər/	'mud'	1	0
dribbel	/'drɪbəl/	'dribble'	1	0
dubbel	/'dʏbəl/	'double'	6	0
fiber	/'fibər/	'fibre'	2	0
flexibel	/flɛk'sibəl/	'flexible'	2	0
hebben	/'hɛbən/	'to have'	30	0
kabbel	/'kɑbəl/	'lap'	1	0
kabel	/'kabəl/	'cable'	1	0
kwebbel	/'kʋɛbəl/	'chatter'	1	0
lobbes	/'lɔbəs/	'good-natured dog'	1	0
oktober	/ɔk'tobər/	'October'	2	0

H Classifications of the alveolar geminate stops

General remarks:

- The positions of the main stresses are indicated, both in content words and in function words. These main stresses are not always realized.
- Unless indicated otherwise, the verb is in a singular form.
- The string "-s" denotes the verbal suffix third person singular present tense.

Table H.1 Numbers of voiced and voiceless geminates in word-combinations ending in daar /dar/ *'there'. The numbers are broken down by the type of the first word of the combinations.*

Type of first word			Numbers of geminates	
			voiced	voiceless
altijd	/ˈɑltɛid/	'always'	0	2
criminaliteit	/kriminaliˈtɛit/	'criminality'	0	1
dat	/ˈdɑt/	'that'	1	2
gehad	/xəˈhɑd/	'had' perfect participle	0	1
goed	/ˈxud/	'good'	0	1
het	/ət/	'it'	1	1
laat	/ˈlat/	'leave'	0	1
met	/ˈmɛt/	'with'	0	1
overheid	/ˈovərhɛid/	'government'	0	1
wat	/ˈʋɑt/	'what'	0	4
zat	/ˈzɑt/	'sat'	1	0
ziet	/ˈzi-t/	'sees'	0	1
zit	/ˈzɪt-t/	'sits'	0	1

Table H.2 Numbers of voiced and voiceless geminates in word-combinations ending in dan /dɑn/ *'then'. The numbers are broken down by the type of the first word of the combinations.*

Type of first word			Numbers of geminates	
			voiced	voiceless
besluit	/bə'slœyt/	'decide'	0	1
dat	/'dɑt/	'that'	1	12
doet	/'du-t/	'does'	0	1
gaat	/'xa-t/	'goes'	0	10
had	/'hɑd/	'had'	0	2
het	/ət/	'it'	1	7
inrijdt	/'ɪnrɛid-t/	'drives into'	0	1
kwijt	/'kʋɛit/	'lost'	0	1
met	/'mɛt/	'with'	0	1
moet	/'mut-t/	'must'	0	1
pakket	/pɑ'kɛt/	'parcel'	0	2
praat	/'prat/	'talk'	0	1
staat	/'sta-t/	'stands'	0	2
tijd	/'tɛid/	'time'	0	3
vooruit	/vor'œyt/	'all right'	0	1
wat	/'ʋɑt/	'what'	0	10
weet	/'ʋet/	'know'	0	1
we't	/'ʋət/ < /ʋə ət/	'we it'	0	2
zat	/'zɑt/	'sat'	0	1
ziet	/'zi-t/	'sees'	0	2
zit	/'zɪt(-t)/	'sit(s)'	0	4

Table H.3 Numbers of voiced and voiceless geminates in word-combinations ending in the conjunction, the neuter singular demonstrative, or the relative pronoun dat /dɑt/. *The numbers are broken down by the type of the first word of the combinations.*

Type of first word			Numbers of geminates	
			voiced	voiceless
afscheid	/ˈɑfsxɛid/	‘farewell’	0	1
altijd	/ˈɑltɛid/	‘always’	2	1
dat	/ˈdɑt/	‘that’	4	12
deed	/ˈded/	‘did’	1	0
doet	/ˈdu-t/	‘does’	2	4
feit	/ˈfɛit/	‘fact’	3	8
gebied	/xəˈbid/	‘area’	1	1
gehad	/xəˈhɑd/	‘had’ perfect participle	1	6
gezet	/xəˈzɛt/	‘set’ perfect participle	0	1
giet	/ˈxit-t/	‘pours’	0	1
god	/ˈxɔd/	‘god’	1	0
goed	/ˈxud/	‘good’	1	5
groot	/ˈxrot/	‘large’	1	0
had	/ˈhɑd/	‘had’	0	1
heet	/ˈhet-t/	‘is called’	0	4
inderdaad	/ɪndərˈdad/	‘indeed’	0	2
met	/ˈmɛt/	‘with’	0	5
moet	/ˈmut/	‘must’	0	5
niet	/ˈnit/	‘not’	5	9
omdat	/ɔmˈdɑt/	‘therefore’	2	1
pakket	/pɑˈkɛt/	‘parcel’	0	1
tijd	/ˈtɛid/	‘time’	1	1
totdat	/tɔtˈdɑt/	‘until’	0	1
wat	/ˈʋɑt/	‘what’	1	4
weet	/ˈʋet(-t)/	‘know(s)’	0	8
ziet	/ˈzi-t/	‘sees’	1	3
zit	/ˈzɪt-t/	‘sits’	0	1

Table H.4 Numbers of voiced and voiceless geminates in word-combinations ending in the determiner de /də/. *The numbers are* broken down by the type of the first word of the combinations.

Type of first word			Numbers of geminates	
			voiced	voiceless
altijd	/'ɑltɛid/	'always'	0	2
dat	/'dɑt/	'that'	0	13
doet	/'du-t/	'does'	0	2
gaat	/'xa-t/	'goes'	0	1
groot	/'xrot/	'large'	0	1
het	/ət/	'it'	1	1
houdt	/'hɑud-t/	'keeps'	0	1
laat	/'lat/	'leave'	0	1
met	/'mɛt/	'with'	5	25
moet	/'mut-t/	'must'	0	1
nadat	/na'dɑt/	'after'	0	1
net	/'nɛt/	'just'	0	1
niet	/'nit/	'not'	1	2
omdat	/ɔm'dɑt/	'because'	1	1
straat	/'strat/	'street'	0	1
tot	/'tɔt/	'until'	0	1
uit	/'œyt/	'out of'	1	14
voordat	/'vordɑt/	'before'	1	0
wat	/'ʋɑt/	'what'	0	3
weet	/'ʋet/	'know'	0	1
zit	/'zɪt(-t)/	'sit(s)'	1	1

Table H.5 Numbers of voiced and voiceless geminates in word-combinations ending in a form of the verb denken *'to think'. The numbers are broken down by combination type*

Type of word-combination			Numbers of geminates	
			voiced	voiceless
aanbiedt denk	/'anbid-t 'dɛŋk/	'offers think'	1	0
dat denk	/'dɑt 'dɛŋk/	'that think'	7	0
dat dacht	/'dɑt 'dɑxt/	'that thought'	1	0
goed denk	/'xud 'dɛŋk/	'good think'	1	0
het denk	/ət 'dɛŋk/	'it think'	1	0
kwijt denk	/'kwɛit 'dɛŋk/	'lost think'	1	0
uit denk	/'œyt 'dɛŋk/	'out think'	1	0
wat denk	/'ʋɑt 'dɛŋk/	'what think'	1	0

Table H.6 Numbers of voiced and voiceless geminates in word-combinations ending in die */di/ with a meaning different from "he" (cf. §10.2.1). The numbers are broken down by the type of the first word of the combinations.*

Type of first word			Numbers of geminates	
			voiced	voiceless
dat	/'dɑt/	'that'	2	14
gehad	/xə'hɑd/	'had' perfect participle	0	1
had	/'hɑd/	'had' past tense	0	2
heet	/'het-t/	'is called'	0	5
het	/ət/	'it'	1	0
laat	/'lat/	'leave'	0	1
met	/'mɛt/	'with'	8	30
moet	/'mut-t/	'must'	0	3
niet	/'nit/	'not'	1	0
raad	/'rad/	'guess'	0	1
staat	/'stat-t/	'stands'	0	2
uit	/'œyt/	'out of'	0	1
wat	/'ʋɑt/	'what'	0	2
weet	/'ʋet/	'know'	1	0
woordenschat	/'ʋordəsxɑt/	'lexicon'	0	1
ziet	/'zi-t/	'sees'	0	1

Table H.7 Numbers of voiced and voiceless geminates in word-combinations ending in die */di/ with the meaning "he". The numbers are broken down by the type of the first word of the combinations.*

Type of first word			Numbers of geminates	
			voiced	voiceless
deed	/'ded/	'did'	1	2
doet	/'du-t/	'does'	0	5
doordat	/dor'dɑt/	'through the fact that'	0	1
gaat	/'xa-t/	'goes'	0	2
had	/'hɑd/	'had'	1	5
heet	/'het-t/	'is named'	0	1
laat	/'lat-t/	'let' present tense	0	3
liet	/'lit/	'let' past tense	0	1
moet	/'mut-t/	'must'	0	2
omdat	/ɔm'dɑt/	'because'	0	2
staat	/'sta-t/	'stands'	0	2
voordat	/'vordɑt/	'before'	0	1
wat	/'ʋɑt/	'what'	0	6
weet	/'ʋet-t/	'knows'	0	1
zat	/'zɑt/	'sat'	0	1
zit	/'zɪt-t/	'sits'	0	1
zodat	/zo'dɑt/	'so that'	0	1

Table H.8. Numbers of voiced and voiceless geminates in word-combinations ending in a form of ding *'thing'. The numbers are broken down by combination type.*

Type of word-combination			Numbers of geminates	
			voiced	voiceless
dat ding	/'dɑt 'dɪŋ/	'that thing'	5	1
goed ding	/'xud 'dɪŋ/	'good thing'	1	0
het ding	/ət 'dɪŋ/	'the thing'	1	0
met dingen	/'mət 'dɪŋən/	'with things'	2	0
wat dingen	/'ʋɑt 'dɪŋən/	'some things'	2	0

Table H.9 Numbers of voiced and voiceless geminates in word-combinations ending in the proximal demonstrative for singular neuter nouns dit /dɪt/. *The numbers are broken down by the type of the first word of the combinations.*

Type of first word			Numbers of geminates	
			voiced	voiceless
dat	/'dɑt/	'that'	0	2
goed	/'xud/	'good'	1	0
met	/'mɛt/	'with'	1	0
moet	/'mut-t/	'must'	0	1
zit	/'zɪt-t/	'sits'	1	0

Table H.10 Numbers of voiced and voiceless geminates in word-combinations ending in a form of the verb doen *'to do'.* The numbers are broken down by combination type. These types are glossed *literally.*

Combination type			Numbers of geminates	
			voiced	voiceless
aanbod doen	/'anbɔd 'dun/	'make (inf.) an offer'	1	0
dat deed	/'dɑt 'ded/	'that did'	3	0
dat deden	/'dɑt 'dedən/	'that did (pl.)'	1	0
dat doen	/'dɑt 'dun/	'that do (inf. or pl.)'	8	0
dat doe	/'dɑt 'du/	'that do (1st ps. sing.)'	6	0
dat doet	/'dɑt 'dut/	'that does'	6	0
dit doen	/'dɪt 'dun/	'this do (pl.)'	1	0
gaat doen	/'xat 'dun/	'will (3rd ps. sing.) do'	1	0
het deed	/ət 'ded/	'it did'	1	0
het doen	/ət 'dun/	'it do (inf. or pl.)'	3	0
het doet	/ət 'dut/	'it does'	1	1
luistervaardig-heid doen	/lœystər'vardɪx hɛid 'dun/	'listening skill to do'	1	0
moet doen	/'mut 'dun/	'must do'	3	0
niet deden	/'nit 'dedən/	'not did (pl.)'	0	1
niet doen	/'nit 'dun/	'not do (inf. or pl.)'	9	0
niet doet	/'nit 'dut/	'not does'	1	0
niet doe	/'nit 'du/	'not do'	1	0
tijd doe	/'tɛid 'du/	'time do'	1	0
universiteit doet	/yniversi'tɛit 'dut/	'university does'	1	0
wat doen	/'ʋɑt 'dun/	'what do (pl.)'	2	0
woordenschat doen	/'ʋordəsxɑt 'dun/	'lexicon to do'	1	0

Table H.11 Numbers of voiced and voiceless geminates in word-combinations ending in dus /dʏs/ *'therefore'. The numbers are broken down by the type of the first word of the combinations.*

Type of first word			Numbers of geminates	
			voiced	voiceless
bestaat	/bə'sta-t/	'exists'	0	1
bezit	/bə'zɪt-t/	'posseses'	1	0
dat	/'dɑt/	'that'	3	0
dit	/'dɪt/	'this'	1	0
gaat	/'xa-t/	'goes'	2	0
goed	/'xud/	'good'	1	0
had	/'hɑd/	'had'	1	2
het	/ət/	'it'	1	0
moet	/'mut(-t)/	'must'	2	0
wat	/'ʋɑt/	'what'	0	1
weet	/'ʋet/	'know'	0	1
zat	/'zɑt/	'sat'	1	0
ziet	/'zi-t/	'sees'	1	0
zit	/'zɪt/	'sit'	1	0

Table H.12 Numbers of voiced and voiceless geminates in word-combinations ending in door /dor/ *'through' / 'by'. The numbers are broken down by the type of the first word of the combinations.*

Type of first word			Numbers of geminates	
			voiced	voiceless
gaat	/'xa-t/	'goes'	1	0
goed	/'xud/	'good'	1	0
altijd	/'ɑltɛid/	'always'	1	0
niet	/'nit/	'not'	2	0

Table H.13 Numbers of voiced and voiceless geminates in combinations ending in the enclitic form d'r /dər/ *of* daar *'there'. The numbers are broken down by the type of the first word of the combinations. The combinations starting with verb forms or* dat *are listed in Table G.3.*

Type of first word			Numbers of geminates	
			voiced	voiceless
het	/ət/	'it'	0	6
gat	/'xɑt/	'hole'	0	1
hoe 't	< /'hu ət/	'how it'	0	2
niet	/'nit/	'not'	0	1
omdat	/ɔm'dɑt/	'because'	1	5
totdat	/tɔt'dɑt/	'until'	0	1
voorraad	/'vorad/	'supply'	1	0
wat	/'ʋɑt/	'what'	2	21

Table H.14 Numbers of voiced and voiceless geminates in combinations which do not occur in one of the preceding tables, and of which the second word is a content word. The numbers are broken down by combination type. These types are glossed literally, and ordered alphabetically by the type of their second word.

Combination type			Numbers of geminates	
			voiced	voiceless
goed daarom	/'xud 'darɔm/	'good therefore'	1	0
het daarom	/ət 'darɔm/	'it therefore'	1	0
gaat dag	/'xat 'dɑx/	'goes day'	0	1
groot deel	/'xrot 'del/	'large part'	2	1
het deugt	/ət 'døxt/	'it is good'	1	0
niet deugt	/'nit 'døxt/	'not is good'	1	0
met diamanten	/'mɛt dija'mɑntən/	'with diamonds'	1	0
dat diner	/'dɑt di'ne/	'that dinner'	1	0
het doel	/ət 'dul/	'the target'	1	0
het donkergroen	/ət 'dɔŋkərxrun/	'the dark-green'	1	0
het dons	/ət 'dɔns/	'the down'	1	0
met dons	/mɛt 'dɔns/	'with down'	1	0
het dubbele	/ət 'dʏbələ/	'the double'	1	0
het Duits	/ət 'dœyts/	'the German language'	2	0
vanuit Duitsland	/vɑn'œyt 'dœytslɑnd/	'from Germany'	1	0
gaat duren	/'xat 'dyrən/	'will take'	1	0
niet durven	/'nit 'dʏrvən/	'do not dare'	1	0
wat duurder	/'ʋɑt 'dyrdər/	'somewhat more expensive'	1	0
dat duurt	/'dɑt 'dyr-t/	'that takes'	3	1

Table H.15 Numbers of voiced and voiceless geminates in combinations which do not occur in one of the preceding tables, and of which the second word is a function word. The numbers are broken down by combination type. These types are glossed literally, and ordered alphabetically by the type of their second word.

Combination type			Numbers of geminates	
			voiced	voiceless
tijd da'k	/'tɛid 'dɑk/ < /'tɛid 'dɑt 'ɪk/	'time that I'	0	1
weet da'k	/'ʋet 'dɑk/ < /'ʋet 'dɑt 'ɪk/	'know that I'	0	1
met deze	/'mɛt 'dezə/	'with these'	5	1
wat deze	/'ʋɑt 'dezə/	'what these'	1	0
met d'r	/'mɛt dər/	'with her (enclitic)'	0	1

Mirjam Ernestus

Voice Assimilation and Segment Reduction in Casual Dutch

A Corpus-Based Study of the Phonology-Phonetics Interface

This book is the first extensive study of the pronunciation of casual Standard Dutch. It provides completely fresh data, and concludes that most properties that are characteristic of casual, as opposed to careful, Dutch are not necessarily the result of phonological processes. They are the effects of the storage of word-combinations and reduced variants in the lexicon, and, above all, the speaker's natural tendency to reduce articulatory effort.

The study focuses on the realization of vowels as schwas, the perceptual absence of segments, and, above all, the realization of obstruents as voiced or voiceless. The hypothesis is proposed that coda and word-final obstruents are realized as voiced or voiceless depending on which realization requires no additional articulatory effort. This hypothesis is shown to be in accordance with all data from previous studies as well as with fresh data. It is taken as the starting point for a detailed analysis of the realization of obstruents in all word and syllable positions in Dutch.

The data forming the basis of the study are taken from a corpus of spontaneous conversations which was built especially for the purposes of this investigation. The design of the corpus and the processing of the data receive ample attention, in particular the auditory classification of intervocalic stops as voiced or voiceless and the relation between this classification and the acoustic properties of the stops.

This book is of interest to anyone concerned with the pronunciation of Dutch, casual speech, the phonology and phonetics of [voice], or corpus linguistics.

ISBN 90-76864-02-0